

令和元年度厚生労働科学研究費補助金（政策科学総合研究事業）  
研究報告書

児童虐待対応におけるリスクアセスメントのためのデータ収集基盤構築と  
AIを活用したリスク評価に向けた研究に関する研究

研究代表者 高岡 昂太 産業技術総合研究所

研究要旨：児童虐待におけるリスクには、虐待の発生リスク、虐待による死亡（致死的行為の発生）リスク、虐待の再発や継続のリスク、虐待が子どもの心身の健全な発達に悪影響を及ぼすリスク、世代間連鎖のリスクなどが想定される。的確なリスクアセスメントとは、包括的な情報収集と評価によって、これらのリスクを事前に見抜き、それに応じた対応を講じるまでの一連のプロセスを指す。このとき、リスクアセスメントの過程には、未来の出来事を含めた未観測の事態に対する「予測」が含まれている。こういった不確実な状況に対し、十分な説明根拠をもって介入や支援を講じるためには、データに基づく知見の活用が欠かせない。児童相談所を代表とする公的機関として、対応の一貫性を確保する上でも、過去のデータを参照することは不可欠となる。

本研究では、児童虐待対応におけるリスクアセスメントのためのデータ収集基盤と、AIを活用したリスク評価に向けた研究を実施した。データに基づくリスクアセスメントを実現するために必要な情報基盤について検討し、解析技術を用いたリスク評価と、リスクアセスメント情報の他の活用可能性について検討することを目的とした。当該目的を達成するために、9つの個別課題(研究Iから研究X)を設定した。

## 研究I

### A. 研究目的

客観的かつ透明性の高い、的確な対応判断を講じる上で、過去のデータを参照することは欠かせない。研究Iでは、個別の自治体に協力を仰ぎ、各組織のデータベースに蓄積されている項目情報を整理することで、既存のデータ基盤の構造や内容を予備的に把握する。そして、データベースの設計等に関する現状の課題点を整理することで、改善点や理想的な設計のあり方について議論する。

### B. 研究方法

都市部および地方部を含めた5自治体の児童相談所データベースのバックアップファイルまたはcsvエクスポートファイルを取得し、項目内容やテーブル構造について整理した。

#### (倫理面への配慮)

自治体の名称と詳細な項目名を除いた、データベースの概要情報にとどめる配慮を講じた。また、当該研究手続きに関しては、下記の倫理審査手続きにて承認を得ている。

【国立研究開発法人 産業技術総合研究所 人間工学実験委員会 事前申請】

整理番号：人2021-1157(新規)

実験課題名：自治体が取得した児童相談所データベースにAIを活用した効果的な虐待対応手法の研究(2021)

判定結果：人間工学実験審査申請非該当

【社会福祉法人恩賜財団母子愛育会 愛育研究所研究倫理委員会】

通知日：2019年1月15日

研究課題：児童相談所及び市区町村の児童相談データベースを活用したAI応用研究

判定結果：承認(承認番号第4号)

### C. 研究結果

**対象自治体** 国内の都道府県の内、協力の得られた5つの自治体を対象とした。いずれも、管轄地域内に複数の児童相談所を有する都道府県となる。管轄する地域の基礎情報等は、特定可能性を排除するため割愛する。

**データベースの構造と機能** 対象となった5つの自治体について、その全てがリレーショナルデータベースを使用していた。データベースに含まれるテーブル(表)の設計は自治体ごとに異なっており、事例を管理する単位(通告の単位、事例の単位、子どもの単位、家族の単位等)を複数有する場合があるものの、自治体間でシステム上の記録方式に部分的な差異が確認された。

**項目内容** データベースのテーブルに含まれる事

例関連情報には、行政管理上必須となる項目については全ての自治体で共通していたが(ただし、記録上の項目名やフィールドの文言等は異なる)、事例のアセスメントに関する情報については、その記録の有無や方式で大きな違いが認められた(表1)。また、アセスメント関連情報を取得・記録するタイミングや回数、アセスメントの具体的該当情報を全て記録するか判定結果(進行管理上のラベル、「重篤度」や「緊急度」などの評価ラベル)のみを記録するかなど、情報の記録と管理方法にも違いが認められた。なお、子どもの年齢等に関する基礎情報を除き、(リスク)アセスメントに関連する項目の大半は、1自治体を除いて、福祉行政報告例等への報告が求められる基礎情報のみ(相談種別、一時保護、措置内容等)であった。

表1. 各対象自治体の利用データベースの基礎特徴

	組織	アセスメント情報	項目内容	その他基礎情報	措置内容	備考
自治体A	児相	25項目	独自項目	福祉行政報告例	○	通告毎・受理種別変更ごとにアセスメント、短文ケース記録文書情報あり
自治体B	児相	判定結果のみ	独自項目	福祉行政報告例	○	リスク判定結果が随時最新情報に上書き更新、短文ケース記録文書情報あり
自治体C	児相	判定結果のみ	不明	福祉行政報告例	○	リスク判定結果が随時最新情報に上書き更新、短文ケース記録文書情報あり
自治体D	児相	判定結果のみ	不明	福祉行政報告例	○	
自治体E	児相	判定結果のみ	不明	福祉行政報告例	○	

### D. 考察

調査への協力が得られた5つの自治体について、各自治体の児童相談所が利用している児童虐待関連情報を蓄積するデータベースの構造について、基本的な特徴を整理した。その結果、自治体ごとに、データベースに含まれるテーブルの内容や構造が異なっており、アセスメント情報を記録するタイミングや内容についてもいくつかの差異が見受けられた。

特に、どのような単位で事例情報を記録するかについては、蓄積する情報を活用する際に大きく影響する。具体的には、「当該事例で将来的に再度の虐待通告が発生するか」や「どのような事例が将来的な再度の通告に至りやすいか」といった課題を検証する際に、「子ども一人一人」の単位で状態像を理解するか、「家族の単位」で状態像を理解するかなど、全く異なる視点での知見整理になるためである。データ活用を前提とする場合には、様々な活用可能性を考慮して、最大限細かい粒度で情報を蓄積することが望ましいと言える。

また、「情報を記録するタイミング」についても、

明確に区分されていることが望ましい。例えば、「子どもの頭部に外傷がある」などのアセスメント情報があったとして、それが「通告時点」で確認されている情報なのか、「現認」の段階で得られた情報か、「継続的な支援中」に得られた情報なのかで、他の情報との関連を検討する際に意味合いが全く異なるものとなる。「どの時点で得られた情報か」については、各組織の業務フローの在り方やその標準化にも関連する。標準的なデータ蓄積を考えるにあたっては、自治体ごとの業務フローの実態と標準的(部分)の模索等に関する調査等が必要になると考えられる。

また、本調査において、事例の詳細なアセスメントデータを整然な(分析に活用できる)形式で蓄積していたのは一自治体にとどまった。リスクアセスメント情報の有無で再発等の予測精度は変化し、より詳細なアセスメント情報がある場合の精度が高い(厚生労働省, 2019)。データの利活用に係るアセスメントデータの蓄積は、よって最大限詳細かつ分析可能な形式にて実施されることが望ましいと言える。

#### E. 結論

事例内容に関する情報項目は、項目情報を蓄積していた一自治体を除いて、福祉行政報告例での年次報告が求められる次の項目等が中心となっていた。精度の高い各種リスクの予測を実現する上では、リスクアセスメント項目情報の蓄積が必要となる。業務フローなど自治体独自の観点を含めつつ、リスクアセスメント等の基本情報部分は標準的であることが望ましい。また、多面的な活用を想定して、リスクアセスメントデータ用の TABLE を設置することや、利活用を見越したエクスポート機能の実装が標準的に必要であると考えられる。

#### F. 健康危険情報（総括研究報告書にまとめて書きたい?）

特になし

#### G. 研究発表

##### 1. 論文発表

なし

##### 2. 学会発表

なし

#### H. 知的財産権の出願・登録状況

(予定を含む。)

##### 1. 特許取得

なし

##### 2. 実用新案登録

なし

##### 3. その他

なし

#### 【引用文献】

厚生労働省(2019). 平成 30 年度子ども・子育て支援推進調査研究事業(調査研究課題番号 26) 体系的な子ども虐待データベースの構築及びデータに基づくリスクアセスメントの効果に関する調査研究(事業主体: 国立研究開発法人 産業技術総合研究所)。 URL: [https://staff.aist.go.jp/kota.takaoka/Ai%20for%20Obetter%20society\\_files/pdf/report-project26.pdf](https://staff.aist.go.jp/kota.takaoka/Ai%20for%20Obetter%20society_files/pdf/report-project26.pdf) (Retrieved 2021.12.22)

## 研究 II

### A. 研究目的

アセスメントデータを活用した的確なリスク予測を実現するためには、必要な情報が活用可能な形式で蓄積されている必要がある。研究 I では、事例の特徴を反映したリスク項目該当情報を蓄積していた自治体は対象自治体のうち 1 例にとどまり、その他の自治体では基礎項目の該当情報のみの蓄積となっていた。そこで、研究 II では、データの利活用を見据えた蓄積データの内容について検討する。

蓄積するデータの内容については、様々な観点からの検討が可能となる。事例の実態を詳細に記述したテキスト情報や、年次統計への計上・報告内容など、目的に応じて採用が求められる観点は異なる。本研究の文脈に照らした「リスク」に関する視点からは、(1)すでに把握されている虐待状況に関する項目の該当情報、(2)虐待の発生や維持の要因に係る子どもや保護者、家庭環境等の情報、(3)将来的な虐待の発生や重症化に関する予測の情報などが、的確な相談援助活動を展開する上で必要になると考えられる。

本研究では、児童虐待に関するリスクの中でも、アウトカムの重大性から致死的行為の発生リスク(死亡リスク)に焦点を当てて、その予測に貢献するアセスメント項目について検討する。厚生労働省(2021)では、既に重篤な虐待の発生予測に貢献する標準アセスメント項目が提案されている。当該項目を軸として、機械学習を用いたリスク予測モデリングを行い、試験的な予測精度の検証を実施する。なお、重大なアウトカムを予測する際には、「誤予測 (misprediction/prediction error)」に対して最大限の配慮を講じる必要がある。本研究では、機械学習技術の中でも、高い予測性能を持つことから頻繁に活用される Extreme Gradient Boosting(XGBoost)モデルを使用する。予測の誤りを最大限排除することを目指す。また、「予測結果の出力時点で当該予測がどの程度誤りである可能性があるか」については、ベイズ推論の枠組みで構成されたモデルで出力される不確かさ(uncertainty)の指標で評価されることがある(Lee

et al., 2018). そこで、本研究では、ベイズ多層ニューラルネットワークモデルを活用した予測と誤予測の事前評価を行う。

## B. 研究方法

厚生労働省(2021)で提案された標準アセスメント項目群を使用し、実際の事例データを用いて重篤な虐待状況の発生に関する予測精度を検証する。

**使用データと背景調査** 厚生労働省(2020)で実施された、全国調査データを使用する。当該調査は、児童相談所(215箇所)および市区町村虐待対応関連部門(1894箇所)を対象に実施された全国 Web 調査(2019年11月から12月にかけて実施)である。各組織で最大20事例の児童虐待事例について、30のアセスメント項目への該当状況が報告された。提示されたアセスメント項目は、文献レビューによって抽出された420項目から、組織毎に無作為に提示された。なお、アセスメント項目以外にも、事例の基礎情報として、児童および主たる養育者の年齢と性別、主たる虐待種別、主たる虐待者種別、過去の虐待係属歴、DV・面前暴力事案、特定妊婦への該当状況が含まれた。また、重篤な虐待への該当状況を測定する項目として、重篤な身体的虐待、重度ネグレクト、性的虐待(疑いを含む)のそれぞれへの該当情報も調査されている。調査の結果、5493件の児童虐待事例情報が収集された。なお、厚生労働省(2020)では、420項目のアセスメント項目に関する基礎評価を経て、129項目への項目統合が実施されている。そして、厚生労働省(2021)では、129項目へと統合した項目を用いた追加調査を実施し、1447件の事例データをもとに、重篤な虐待の予測に貢献する31項目を抽出し、標準項目リストとして提案を行っている。これらを踏まえ、研究Ⅱでは、厚生労働省(2020)で収集された5493件の事例データから、厚生労働省(2021)で提案された31項目の標準項目を選抜したデータセットを作成し、それを使用した。

**予測対象** 厚生労働省(2020)および厚生労働省(2021)では、重篤な虐待状況を示す区分として、重篤な身体的虐待、重度ネグレクト、性的虐待(の疑い)、その他の重篤な虐待の4種が設定されている。特に、重篤な身体的虐待と重度ネグレクトに関しては、子どもの生命の危機に直接係る虐待行為となることから、これら2つの重篤な虐待を予測対象として設定した。重篤な身体的虐待および重度ネグレクトに関する個別内容については、表2.1のいずれか一つ以上に該当した場合に「該当」と判断される項目である。

表 2.1 重篤な身体的虐待と重度ネグレクトの定義リスト(厚生労働省(2021)より引用)

区分	No	重篤な虐待の記述的定義(項目内容)
重篤な 身体的虐待	1	頭蓋内損傷・頭蓋内出血がある(不慮の事故によるもの確証がないもの)
	2	眼底出血・網膜剥離・水晶体脱臼などの眼科所見がある(不慮の事故によるもの確証がないもの)
	3	内臓損傷・腫脹・出血がある(不慮の事故によるもの確証がないもの)
	4	毒物・薬物等による中毒症状(飲まされるなどによる)がある
	5	胸部・顔面・胸部・腹部の打撲痕がある
	6	骨状軟骨や二重条痕などの特徴的な形状の創傷がある
	7	新旧の多数の創傷・骨折痕がある(不慮の事故によるもの確証がないもの)
	8	身体各部に拘束を疑う傷がある(不慮の事故によるもの確証がないもの)
	9	熱傷、熱傷痕、高熱の前兆があったとみられる熱傷、熱傷痕がある(不慮の事故によるもの確証がないもの)
	10	頸部絞扼(首絞め)の疑わしき痕がある
	11	殴る、蹴る、叩く、踏みつけるなどの暴力によって生じる可能性のある重度の傷・あざ・骨折がある
	12	道具による暴力行為によって生じる可能性のある重度の傷・あざ・骨折がある
	13	児童の安全を守る監督者がいない状態で危険にさらされる戸外にいる(締め出しの疑い)
	14	頸部絞扼(首絞め)の目撃・報告がある
	15	児童を溺れさせる(風呂等に沈める行為を含む)
	16	児童の鼻と口をふさぐ(乳幼児の場合は布を顔にかける行為を含む)
	17	児童を縛りつける、トランク・箱などの閉所に入れられないようにして閉じ込める、児童を布団裏にする
	18	殴る、叩く、蹴る、踏みつける、あるいは道具による暴力行為など、児童が怪我を負う危険性がある
	19	乳幼児を立て抱きにして前後に激しく揺さぶる・投げ飛ばす・床に落とす
重篤な ネグレクト	20	医療によるミュンヘンワグゼン症候群(MSBP)
	21	その他重篤な身体的虐待
	22	身体的な理由にもかかわらず児童の身長または体重が標準身長・標準体重の2SDを下回っている
	23	児童が常態・置きまり、放置(車中含む)されている。または、監護責任者不在での夜間徘徊がある
	24	児童が病弱などにより衰弱している状態のまま放置されている
	25	養育者の監護なしに、児童だけでは危険な環境におかれている
	26	感染症、乳幼児の下痢、発熱疾患、または重度の外傷等があるが病院を受診できていない
	27	養育放棄がある。または養育者が児童の養育(身体的対応含む)に関して無関心・拒否的・「世話をしたくない」などの断えがある
	28	毒物や不適切な薬物を、児童が勝手に飲めしてしまう環境がある
	29	その他重篤と判断するネグレクト

**適格基準と項目抽出** レコード源からの情報抽出に際して、適格基準は定めなかった。また、解析使用データの抽出作業に際しても、除外基準は設定しなかった。アセスメント項目等に欠損がある場合には、データの前処理で補完を実施するため、除外対象としなかった。項目の抽出に関しては、表2.2に示した厚生労働省(2021)による標準項目リストに加えて、児童および主たる養育者の年齢と性別を抽出した。なお、厚生労働省(2020)で収集された解析に使用するデータに含まれない項目については、本解析において補填等は実施しなかった。

表 2.2 重篤な虐待の発生予測に貢献する標準項目リスト(厚生労働省(2021)より引用)

区分	No	重篤な虐待の発生予測に貢献する標準項目内容
外傷	1	養育者の説明の回避、または説明内容に疑念がある
	2	支援や介入の困難、または支援のための資源が不足している
家庭環境	3	養育負担の偏りがある、または、夜間看護がない
	4	養育環境が不適切である
	5	経済不安、または経済的不安定さがある
	6	情緒的な問題、対人距離、または受容関係に課題がある
	7	養育者を過剰に支持する
児童	8	育てにくさがある児童である
	9	養育者に対して挑発やユスカレートする行為がある
	10	摂食や排泄の異常、または喘息やアレルギーがある
	11	過剰なしつけ、体罰での暴力、正座等の強制、または暴力のほめかしがある
養育者	12	養育者の都合が養育より優先している、生活が自己中心的である、態度から事象改善が見込まれない
	13	育児スキルの不足や育児の不履行がある
	14	怒りや発見的事態への対応が困難
	15	支援の拒絶、回避、もしくは無関心さがある、または支援者への態度に一貫性がない
	16	精神科の通院歴、不安定さ、判断力の減退、または養育困難がある
	17	養育者の家庭外ストレスがある
	18	児童に被害歴がある
過去	19	きょうだい相談歴、一時保護歴、または措置歴等がある
外傷	20	爛傷、おすかな傷、または説明されない傷がある
	21	家族構成、または同居人に変化がある
家庭環境	22	家族構成、または同居人に変化がある
	23	児童の身体に打撲痕や内出血などの外傷がある
	24	家庭での虐待を認めている
	25	虐待による傷つきのフラッシュバックを感じている
	26	虐待の歴史、編纂、認識欠如がある
養育者	27	夫婦関係、対立、もしくは話し合いの困難さがある、または夫婦間の立場が対等ではない
	28	支援者への攻撃性がある
	29	若年出産(10代での妊娠・出産)である
	30	人前での子育て能力がある、または思いもよらない機嫌がある
	31	物質や行為への依存がある、または発達障害の診断や疑いがある
	その他	32

**データの事前処理** 解析データに対して、欠損処理、学習・検証データの分割、クラス不均衡の解消の3つの事前処理を実施した。調査構造に基づき無作為に発生するアセスメント項目の該当情報に関する欠損に対しては、0を代入補完した。情報の不足する児童虐待対応の初動を想定し、「該当が確認されない状況」をデータ上で模擬的に表現した。その他、欠損が含まれる養育者の年齢に関しては、データ全体の中央値を代入して補完した。学習・検証データは無作為分割により作成し、構成比を80%(学習データ)と20%(検証データ)とした。なお、重篤な身体的虐待と重度ネグレクトの該当率は、解析データ内でクラス不均衡(非該当が極端に多いなど)となる。不均衡データに由来する予測の誤りを緩和するため、学習データに対してSMOTEサンプリングを適用した(Chawla et al., 2002)。SMOTEサンプリングの適用条件は、統計解析環境RのパッケージDMwRの標準規定値を採用した。

**モデルと学習条件(XGBoost)** 重篤な虐待の発生を予測するモデルとして、XGBoostを使用した。XGBoostとは、勾配ブーストと呼ばれる最適化アルゴリズムを用いた機械学習手法の一種であり、対象が観測されるパターンを条件分岐の形式で捉えるモデルである(Chen, C. & Guestrin, C., 2016)。XGBoostの主要な学習条件(ハイパーパラメータ)として、損失関数は対数損失(logloss)を使用し、学習率(eta)は0.03、条件分岐の最大の深さは(max\_depth)4とした。ブースティングのラウンド数(num\_round)は、交差検証法(5-fold cross validation)により1回から8,000回の範囲で最適な値を決定した。その他の学習条件は、統計解析環境Rのxgboostパッケージにおける標準設定値を利用した(Chen, T. & Guestrin, C., 2016)。

**モデルと学習条件(ベイズ多層ニューラルネットワークモデル)** 誤予測の事前評価を実施するため、

本研究ではベイズ多層ニューラルネットワークによる予測を合わせて実施した。

多層ニューラルネットワークとは、ヒトの神経ネットワークを模した解析モデルの一種である。入力と出力(予測)の間に複数ノードで構成される層(中間層)が設定され、入力に対する重み付けを各層で繰り返しながら、最終的に出力層へと情報が伝達される構造を持つ。そして、出力層での値が予測対象ラベルと整合するように、重み付けの度合いが学習される。本研究で使用したモデルを以下に定義する。

学習に使用する事例データセット $X$ は、実数の $N \times D$ 行列である( $X$ を構成する $N$ 個の行ベクトルは、それぞれ、 $D$ 個の項目からなる一つの事例データに対応する)。また、各事例における重篤な虐待の発生の有無を示す予測対象ラベル $Y$ は、0または1を成分とする $N$ 次の列ベクトルである。 $Y$ の第 $i$ 成分 $Y_i$ の値が、 $X$ の第 $i$ 行ベクトル $X_i$ の表す事例データにおける重篤な虐待の発生の有(1)無(0)を表す。

本モデルで設定した中間層は3つある。各層に含まれるノード数をそれぞれ $D_1, D_2, D_3$ とする。また、 $D_0$ は入力層の次元 $D$ を表し、 $D_4$ は出力層の次元1を表すこととする。入力層から出力層までの重み付け係数を、それぞれ $W^0, W^1, W^2, W^3$ とする。各 $W^k$ は、実数 $D_k \times D_{k+1}$ 行列である。各層における共通のバイアスを、実数 $b^0, b^1, b^2, b^3$ とする。

モデルの活性化関数 $f(z)$ は、次式で定義した:

$$f(z) = \frac{1}{1 + e^{-z}} \quad (2.1)$$

第一層から第三層までの中間層の各ノードで保持される値の組を、 $V^1, V^2, V^3$ と表す。また、 $V^0$ は入力層に与えられるデータセット $X$ を表すこととし、 $V^4$ は出力層で保持される値の組み $\theta$ を表すこととする。各 $V^k$ は、実数の $N \times D_k$ 行列である。 $\theta$ は、 $N \times 1$ 行列、すなわち $N$ 次の列ベクトルで、各成分 $\theta_i$ に対して、 $0 < \theta_i < 1$ が成り立つ。各 $V^k$ の $(i, j)$ 成分 $V_{i,j}^k$ の値は、事例データセット $X$ の第 $i$ 行ベクトル $X_i$ と、重み付け係数 $W^k$ の第 $j$ 列ベクトル $W_j^k$ を用いて、以下のように定義される:

$$V_{i,j}^{k+1} = f(b^k + V_i^k W_j^k), \quad k = 0, 1, 2, 3, \quad i = 1, 2, \dots, N, \quad j = 1, 2, \dots, D_{k+1} \quad (2.2)$$

モデルパラメータの推定にはハミルトニアンモンテカルロ法を使用した。各種添字を $i = 1, 2, \dots, N$ ,  $k = 0, 1, 2, 3$ ,  $l = 1, 2, \dots, D_k$ ,  $m = 1, 2, \dots, D_{k+1}$ としたとき、事前分布は、次のように仮定した:

$$\begin{aligned} Y_i &\sim \text{Bernoulli}(\theta_i) \\ b^k &\sim \text{Normal}(0, 1) \\ W_{l,m}^k &\sim \text{Normal}(0, 1) \end{aligned}$$

### (2.3)

なお、本研究では各中間層のノード数を全て 36 に固定した( $D_1 = D_2 = D_3 = 36$ )。パラメータ推定にかかるサンプリング条件は、iteration = 500, warmup = 250, chains = 4 とした(total = 1,000 sample)。

事前の誤予測を評価する際には、 $\theta_i$ の事後分布の幅の広さ(95%確信区間の上限と下限)を使用した。本研究での誤予測の定義は、「重篤な虐待に該当すると予測した時に該当がなかった場合、または、重篤な虐待に該当しないと予測した時に該当があった場合のいずれかに該当する場合」とした。

**性能評価** 予測性能を評価する指標として、PR 曲線下面積(Area Under the Precision-Recall Curve: AUCPR)を採用した。AUCPR は、横軸に感度(Recall)、縦軸に適合率(Precision)を取り、機械学習が出力する予測スコアの判別閾値を変化させた時に現れる曲線(PR 曲線)で区切られた範囲下の面積を用いた指標である。本研究では、PR 曲線下面積を積分によって求めた値を報告する。本研究の文脈において、感度とは「実際に重篤な虐待のあった事例のうち、『重篤である』と予測できた事例の割合」であり、適合率とは、「『重篤である』と予測された事例のうち、実際に重篤な虐待が含まれた割合」を示す。AUCPR、感度、適合率のいずれも 0 から 1 までの値を取り、値が大きいほど高い予測性能を有することを示す。AUCPR は、クラス不均衡なデータセットでも性能評価にバイアスが混入しにくいことから、機械学習研究における報告指標として利用が推奨されている(Sofaer et al., 2019)。また、機械学習による予測と実際が整合していた割合を示す指標として、精度(Accuracy)も補足的に報告する。Accuracy は 0 から 1 までの値をとり、1 に近いほど性能が良いことを示す。クラス不均衡な場合に高い値を取りやすいという性質を持つものの、「正解率」を指し示すものであることから解釈性に優れている。

(倫理面への配慮)

研究 II では、厚生労働省(2020)で実施された児童相談所・市区町村を対象とする全国調査データの二次利用を実施した。調査事業終了後の 5 年間のデータ保存と二次利用については、同意手続きを含めて下記の倫理審査で承認を受けている。

【国立研究開発法人 産業技術総合研究所 人間工学実験委員会 事前申請】

整理番号：人 2019-982 (新規)

実験課題名：児童虐待対応にかかるリスクアセスメント項目の開発と評価

判定結果：人間工学実験審査申請非該当

【社会福祉法人恩賜財団母子愛育会 愛育研究所

研究倫理委員会】

受付番号：第 2 号

研究課題：児童虐待対応にかかるリスクアセスメント項目の開発と評価

判定結果：承認(承認番号第 2 号)

### C. 研究結果

**組入データ** データ抽出の結果、5493 件の事例データが解析に組み入れられた。学習データと検証データをそれぞれ 80%、20%で無作為に分割した結果、重度ネグレクトの二値分類課題では学習データ 4394 件、検証データが 1099 件となった。それぞれに含まれたアウトカムの件数と該当率は、490 件(11.2%)と 126 件(11.5%)であった。クラス不均衡を緩和するため、学習データに SMOTE サンプリングを適用した結果、レコード数は 6370 件、アウトカムの該当率は 23.1%となった。重篤な身体的虐待の二値分類課題では、学習データ 4395 件、検証データが 1098 件となった。それぞれに含まれたアウトカムの件数と該当率は、633 件(11.2%)と 171 件(11.5%)であった。クラス不均衡を緩和するため、学習データに SMOTE サンプリングを適用した結果、レコード数は 8229 件、アウトカムの該当率は 23.1%となった。学習/検証に使用したデータセットの基礎情報を表 2.3 に示す。

表 2.3 予測対象別での学習/検証データセットの基礎情報

key	trainY	trainN	SmoteTrainY	SmoteTrainN	testY	testN	train_ratio	train_smote_ratio	test_ratio
Neglect	490	4394	1470	6370	126	1099	0.1115157	0.2307692	0.1146497
PhysicalAbuse	633	4395	1899	8229	171	1098	0.1440273	0.2307692	0.1557377

**性能評価** 検証データによる予測性能の評価を実施した。XGBoostによる分類性能は、重度ネグレクトの予測で AUC-PR = 0.548、閾値を 0.5 とした場合の精度が 0.903 となった。重篤な身体的虐待の予測では、AUC-PR = 0.551、閾値を 0.5 とした場合の精度は 0.865 となった。一方、ベイズ多層ニューラルネットワークによる分類性能は、重度ネグレクトの予測で AUC-PR = 0.421、閾値を 0.5 とした場合の精度が 0.897 となった。重篤な身体的虐待の予測では、AUC-PR = 0.361、閾値を 0.5 とした場合の精度は 0.850 となった。XGBoost と比較して、AUC-PR が低く、予測性能が劣る傾向が観察された。PR 曲線を図 2.1 に、数値結果を表 2.4 に示す。

表 2.4 XGBoost とベイズ多層ニューラルネットワーク(BDNN)による予測結果

Model	Outcome	AUCPR	Threshold	Accuracy	Precision	Recall	NegPredValue
01_XGBoost	01_Severe Neglect	0.5487001	0.5	0.9035487	0.7631579	0.2301587	0.9085768
01_XGBoost	02_Severe Physical Abuse	0.5511956	0.5	0.8652095	0.8709677	0.1578947	0.8650422
02_BDNN	01_Severe Neglect	0.4216009	0.5	0.8971793	0.6326531	0.2460317	0.9095238
02_BDNN	02_Severe Physical Abuse	0.3613335	0.5	0.8506375	0.5686275	0.1695906	0.8643744

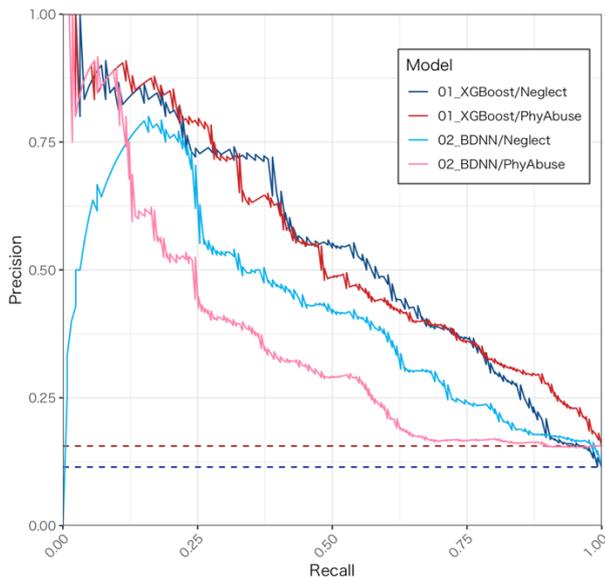


図 2.1 PR 曲線: XGBoost とベイズ多層ニューラルネットワーク(BDNN)による  
重度ネグレクトと重篤な身体的虐待の予測

**誤予測の事前評価** ベイズ多層ニューラルネットワークモデルによって、検証データで算出した $\theta_i$ の事後分布について、その 95%確信区間の幅の大きさと誤予測率との関係を図 2.2 に示す。その結果、重度ネグレクトの分類課題では、事後分布の確信区間の幅が広いほど、誤予測の発生率が高くなる傾向が観察された。他方、重篤な身体的虐待の分類課題では、幅の広さ別で誤予測率が算出されたものの、区間の広さと誤予測率に明確な関連性は観察されなかった。

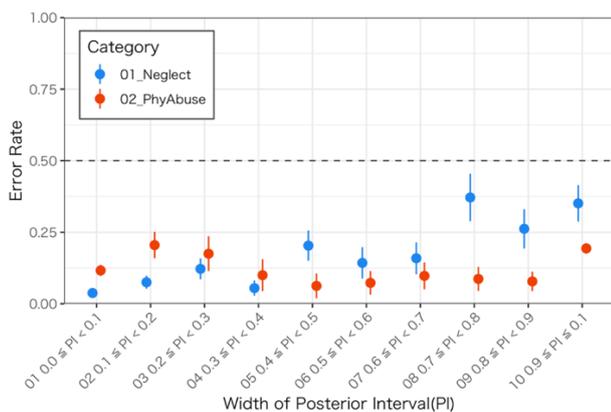


図 2.2 出力層パラメータ( $\theta_i$ )の事後分布の 95%確信区間の幅と誤予測の関係

#### D. 考察

解析の結果、標準アセスメント項目リストを用いて重篤な虐待が一定精度で予測された。一般に、安定して高い性能を発揮するとされる XGBoost モデルには劣るものの、ベイズ多層ニューラルネットワークを用いて、誤予測の事前評価を含めたフレームワークの適用結果が得られた。「誤予測の可

能性をあらかじめ知る」ことができるという点について、その実用性は高いと考えられる。よって、ベイズ多層ニューラルネットワークモデルについては、ノード数や活性化関数の工夫によって精度の向上が必要となる。

なお、いずれのモデルであっても、「重篤な虐待の発生は見逃すことができない」という水準からすると、実用に耐える精度とはならなかった。このとき、本研究で使用したデータは 420 項目から無作為に選抜された 30 項目のみの評定データのみを使用した単発の調査情報であった。すなわち、データに含まれた該当/非該当情報の大半が損失しており、重篤な虐待が確認された段階等の情報にも不足があった。実践利用での予測精度については、自治体等データベースで継続的に蓄積された正確な情報を活用し、厳密な予測対象の定義や欠損発生の可能性を可能な限り排除することで、より向上した性能が得られることが期待される。自治体で蓄積されたデータベース等の情報を活用した検証が求められる。

#### E. 結論

標準的なアセスメント項目情報を活用することで、情報の比較的少ない初動段階であっても、重篤な虐待の将来的な発生や並存が一定精度で予測可能であることが明らかとなった。後続の詳細な検討が必要であるものの、自治体等に継続的データが蓄積されることで、予測性能の向上が見込まれる。

#### F. 健康危険情報 特になし

#### G. 研究発表

1. 論文発表  
なし
2. 学会発表  
なし

#### H. 知的財産権の出願・登録状況 (予定を含む。)

1. 特許取得  
なし
2. 実用新案登録  
なし
3. その他  
なし

#### 【引用文献】

厚生労働省(2020). 令和元年度 子ども・子育て支援推進調査研究事業 事業報告書「課題番号 20: 児童虐待対応におけるアセスメントの在り方に関する調査研究」(事業主体: 国立研究開発法人産業技術総合研究所). URL:

[https://staff.aist.go.jp/kota.takaoka/Ai%20for%20Obetter%20society\\_files/pdf/2020project20-report.pdf](https://staff.aist.go.jp/kota.takaoka/Ai%20for%20Obetter%20society_files/pdf/2020project20-report.pdf) (Retrieved 2021. 05.09.)

Chen, T. & Guestrin, C. (2016). "XGBoost: A Scalable Tree Boosting System", 22nd SIGKDD Conference on Knowledge Discovery and Data Mining. <https://arxiv.org/abs/1603.02754> (Retrieved 2021. 05. 09).

Sofaer HR, Hoeting JA, Jarnevič CS. (2019). The area under the precision-recall curve as a performance metric for rare binary events. *Methods in Ecology and Evolution* 10: 565– 577. doi:10.1111/2041-210X.13140

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321-357.

Lee, J., Bahri, Y., Novak, R., Schoenholz, S., Pennington, J., & Sohl-Dickstein, J. (2018). Deep Neural Networks as Gaussian Processes. *ArXiv*, abs/1711.00165.

### 研究Ⅲ

#### A. 研究目的

児童虐待対応における介入や支援の主たる目的の一つに、再虐待の防止がある。また、リスクアセスメントの評価対象それ自身が再虐待の可能性を予測することに設定される場合もある(加藤, 2001)。一時保護の解除場面等を含め、在宅支援の決定を下す際には、児童の継続的な安全が担保される必要がある。しかし、対応にあたる眼前の事例について考えるとき、当該事例が将来的に再発するか否かを的確に見抜くのは容易ではない。本邦においては、再虐待の発生に関連する事例の特徴に関する研究が進められているものの(Horikawa et al., 2016)、個別事例に対する予測を実現しようとする取り組みは少ない。

本研究では、リスクアセスメントデータと事例の基礎情報から、各事例が将来的に再度通告されるか否か(以下、当該定義を「再発」とする)を、機械学習技術で予測し、予備的な性能評価を行う。研究の文脈から、標準アセスメントツールの項目を用いた再発予測精度の検討が理想となるが、当該項目による縦断的なデータの蓄積はなされていない。そこで、本研究では、リスクアセスメント項目該当情報の経年蓄積がなされている自治体データベース情報を活用して、得られた知見を報告する。

#### B. 研究方法

**対象データ源とリスクアセスメント項目** 解析に使用するデータは、三重県児童相談所のデータベースから抽出した。同県には児童相談所が6箇所あり(うち、一箇所は2019年度新設)、各児童相談所での対応情報が一つのデータベース内に記録される仕組みとなっている。そして、全ての児童相談所にて、2014年より、県独自に作成したリスクアセスメント項目への該当情報の電子記録が開始されている。リスクアセスメントは、通告による相談受付の単位で実施されている。実施のタイミングは、緊急受理会議時点と、一時保護判断時点があり、それぞれの段階でアセスメント項目への該当・非該当・不明が判断される。アセスメント項目は、主に緊急出動を検討する6項目と、一時保護の実施を検討する15項目、合計21項目から構成されている。

**データ源からのレコード抽出手続き** データ源からのレコード抽出は、児童相談所データベースに搭載されたリレーショナルデータベースマネジメントシステムから、リスクアセスメント関連テーブルのエクスポート機能を利用して実施した。抽出の対象期間は、2014年4月1日から、研究実施時点最新月の2021年12月09日までとした。抽出対象は、受付の単位を1レコードとするリスクアセスメントテーブルとし、対象期間中に受理のあった児童虐待通告事例を抽出した。したがって、データ源からのレコード抽出に関する適格基準は、(1)指定期間内での受理であること、(2)児童虐待通告事例であることの2つとした。

**解析データの抽出手続き** データ源から抽出された指定期間内の虐待通告事例から、さらに解析に使用するデータ抽出を行なった。解析データの抽出に際して、虐待種別、児童の年齢に欠損がないことを適格基準とした。また、再発の定義(後述)から、再発ラベルの付与処理後に2020年12月09日以降の事例を除外した。

**項目抽出とデータの前処理** 本研究では、再発の定義を「通告の受付日から起算して、1年(365日)以内に再度の通告があった場合(当該事例が終了していたかは問わない)」とした。同一の児童IDが反復出現したレコードを再発と定義し、再発ラベルを有(1)無(0)で付与した。欠損値について、連続量の場合は中央値代入を実施し、離散量の場合は欠損ラベルを付与した上でOne-hot-encodingによるダミー変数に展開した。リスクアセスメントテーブルから、最終的に解析に組み入れた項目を表3.1に示す。なお、リスクアセスメント情報の記録時点は、「一時保護判断時点」の情報を使用した。

表 3.1 組み入れ項目

変数	フィールド	処理
受付区分	新規・再通告・再受付	one-hot-encoding
主たる虐待者	実父・実母	one-hot-encoding
	実父以外の父・実母以外の母	
	その他	
主たる虐待種別	身体的虐待・性的虐待	one-hot-encoding
	心理的虐待・ネグレクト	
児童性別	男(0)・女(1)	
児童年齢	0~18	正規化
保護者性別	男(0)・女(1)	
保護者年齢	0以上	正規化
アセスメント項目1~21	該当・非該当・不明	label encoding (非該当と不明を0)
一時保護の実施	実施(0)・実施(1)	

**学習/検証データの分割** 学習データと検証データは、80%と 20%の比率で無作為に分割して作成した。

**モデルと学習条件(バイズ多層ニューラルネットワーク)** 個別事例に対する再発予測と、予測時にそれが誤予測である可能性をあらかじめ評価するために、バイズ多層ニューラルネットワークモデルを適用した。解析に使用したモデルの構造は、研究Ⅱと同様である。本研究では、各中間層のノード数を全て 36 に固定した( $D_1 = D_2 = D_3 = 36$ )。パラメータ推定にかかるサンプリング条件は、iteration = 500, warmup = 250, chains = 4 とした (total = 1,000 sample)。事前の誤予測を評価する際には、 $\theta_i$ の事後分布の幅の広さ(95%確信区間の上限と下限)を使用した。

**性能評価** 予測性能を評価する指標として、研究Ⅱと同様に、PR 曲線下面積 (Area Under the Precision-Recall Curve: AUCPR)を採用した。本研究では、PR 曲線下面積を積分によって求めた値を報告する。また、機械学習による予測と実際に整合していた割合を示す指標として、精度 (Accuracy)も補足的に報告する。

(倫理面への配慮)

研究Ⅲは、人を直接の対象として協力や介入等を実施しないデータベース型研究である。当該研究に係る倫理審査については、下記の倫理委員会での審査を経て、承認を得ている。

【国立研究開発法人 産業技術総合研究所 人間工学実験委員会 事前申請】

実験課題名：児童相談所及び市区町村の児童相談データベースを活用した AI 応用研究

判定結果：人間工学実験審査申請非該当(2020年3月9日)

【社会福祉法人恩賜財団母子愛育会 愛育研究所

研究倫理委員会】

受付番号：第 4 号

研究課題：児童相談所及び市区町村の児童相談データベースを活用した AI 応用研究

判定結果：承認(承認番号第 4 号)

### C. 研究結果

**組入データ** データ抽出の結果、15771 件の事例データが解析に組み入れられた。学習データと検証データをそれぞれ 80%, 20%で無作為に分割した結果、学習データ 12616 件、検証データが 3155 件となった。それぞれに含まれたアウトカムの件数と該当率は、4578 件(36.3%)と 1133(35.9%)であった。

**性能評価** バイズ多層ニューラルネットワークによる再発予測の結果、検証データを用いた際の AUC-PR は 0.512 の値が得られた(図 3.1)。



図 3.1 バイズ多層ニューラルネットワークによる再発予測(AUC-PR)

また、検証データに対する予測を実施した際の予測スコアについて、出力層のパラメータ $\theta_i$ に関する事後平均(EAP)の度数分布を正解ラベルごとに示す(図 3.2)。

不明を考慮したベイズ深層学習による予測スコア  
破線は閾値0.5を示す

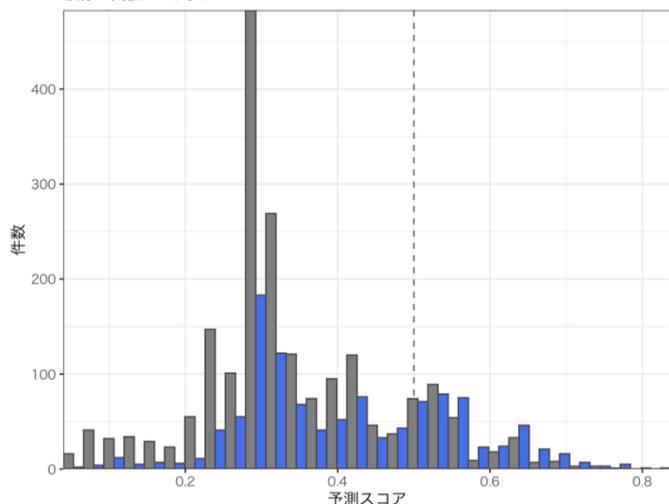


図 3.2 検証データを用いた出力層のパラメータ $\theta_i$ の事後平均の正解ラベル別ヒストグラム  
(青のヒストグラムが実際に再発のあった事例を示す)

出力層のパラメータ $\theta_i$ に関する EAP の値を用いて、閾値を 0.5 に設定し、0.5 より大きい場合を「再発」、0.5 以下の場合を「再発なし」として予測ラベルを付与した。当該条件における各種性能指標は、Accuracy = 0.66 となった。結果の理解可能性を高めるため、実際の再発の有無と予測ラベルのクロス集計(混同行列)を図 3.3 に示す。

ベイズ深層学習による再発予測精度  
判別閾値0.5とした予測と実際のクロス集計(混同行列)  
検証データ3155件でのAccuracy = 0.66

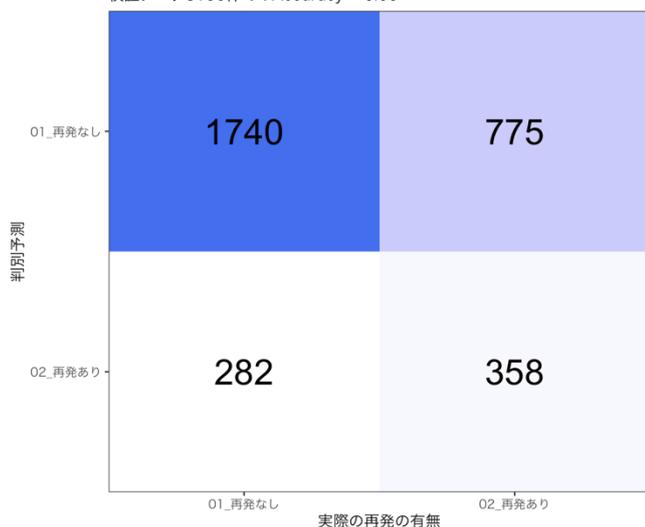


図 3.3 ベイズ多層ニューラルネットワークを用いた  
再発予測の混同行列(検証データで算出)

検証データを用いた再発予測に際して、出力層のパラメータ $\theta_i$ の 95%確信区間の幅の広さと、誤予測との関係を集計した。その結果、事後分布の幅の広さが大きいほど、誤予測率が高くなる傾向

が観察された(図 3.4)。

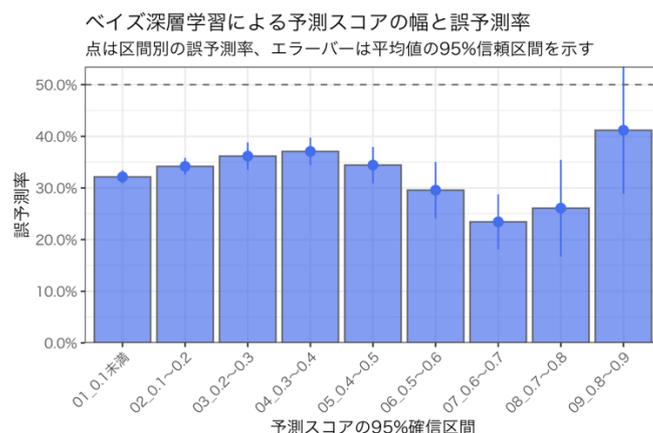


図 3.4 ベイズ多層ニューラルネットワークを用いた  
出力層パラメータの事後分布の 95%確信区間と誤予測率の関係

#### D. 考察

本研究では、一時保護判断時点のリスク情報を用いて、受理後 1 年以内の再発可能性を予測した。その結果、一定程度の予測性能が得られ、一時保護判断時点等の比較的初期段階から事例の再発可能性を評価できる可能性が示された。このとき、再発の可能性を考慮した一時保護判断や在宅復帰判断は、子どもや家庭にとって極めて重大な影響を及ぼしうる問題である。慎重な決断が求められる場面において、機械学習の予測の誤りがもたらす影響は大きい。本研究では、再発予測に係る誤予測の可能性について、予測の出力時点で一定程度評価できることが示された。ただし、総合的な予測精度や誤予測の可能性を評価する上での数値上の精度は十分であるとはならない。今後の精度向上の取り組みや、誤予測の可能性を考慮した確信度の高い予測結果を判断に取り入れてゆくことで、機械学習の予測精度だけに依存せず、総合的に的確な判断が実現される可能性が示された。

#### E. 結論

リスクアセスメント情報を含んだデータベースへの記録情報を活用することで、一定以上の精度で再虐待の可能性が予測できることが示された。こういった再発の可能性は、児童虐待対応における比較的初期段階(受理前後の一時保護判断時点)から評価できる。また、ベイズ多層ニューラルネットワーク等の技術を活用することにより、誤予測の可能性もあらかじめ評価できることが示された。一時保護解除判断や、在宅支援を決定する際の、総合的な判断の中に機械学習等による予測情報を組み入れることで、よりの確かな意思決定が実現される可能性が示唆された。

## F. 健康危険情報

特になし

## G. 研究発表

### 1. 論文発表

なし

### 2. 学会発表

なし

## H. 知的財産権の出願・登録状況（予定を含む。）

### 1. 特許取得

なし

### 2. 実用新案登録

なし

### 3.その他

なし

## 【引用文献】

加藤曜子(2001). 児童虐待リスクアセスメント, 中央法規出版.

Horikawa, H., Sugimoto, S.P., Musumari, M.P., Techasrivichien, T., Ono-Kihara, M. & Kihara, M. (2016). Development of a prediction model for child maltreatment recurrence in Japan: A historical cohort study using data from a Child Guidance Center. *Child Abuse & Neglect*, 59, 55-65.

## 研究IV

### A. 研究目的

本邦における児童虐待通告件数は、年々増加の一途を辿っている。平成 30(2018)年度の児童相談所における児童虐待相談対応件数(速報値)15 万 9838 件となり(厚生労働省, 2020), 30 年連続での件数増加を記録した。増加し続ける児童虐待相談対応件数に対して、限られた支援資源の中で対応の水準を維持するためには、判断の一貫性を担保する確かなマネジメントが必要となる。なかでも、一時保護の実施に係るマネジメントは、児童の安全を確保する上で重大な機能を担う。しかし、一時保護所の満床率は従来から指摘される課題となっている(厚生労働省, 2019)。このような状況下で、年間等を通じて旧来の判断水準を維持することは容易ではない。

一時保護所の定員枠等、支援資源や環境の制約については、本来「児童の最善の利益を追求する」という児童虐待対応の理念に照らせば、必要な資源が常に確保されている必要がある。しかし、増加する相談対応件数を鑑みれば、実際的には、資源の制約に影響を受けて判断水準が経時的に変化し

ていると考えるのが自然である。こういった判断水準の変化を常時把握し、水準を維持するためには、過去のデータを参照したフィードバックが必要になる。

本研究では、リスクアセスメントデータを使用し、一時保護の実施パターン抽出と、新規事例に対する一時保護予測の試験的実践を行う。言い換えれば、「リスクアセスメント情報から、当該事例が過去に一時保護実施されてきたか」に関する予測を個別事例ごとに行い、その判定精度の検証を行う。当該試みは、一貫した一時保護判断の実現や、過去の一時保護判断水準を維持した場合に、どの程度の資源不足(一時保護所の定員や委託枠の超過)が将来的に発生しうるか等の検討にも利用できる可能性がある。

なお、当該課題を検討するにあたり、理想的には標準的なアセスメントデータを用いた保護判断パターンが検出できることが望ましい(自治体間での保護判断の差異などを評価する際にも利用できる)。しかし、現時点では標準アセスメント項目を用いた蓄積データは存在しないため、本研究では、研究Ⅲと同様にして、リスクアセスメント項目該当情報の経年蓄積がなされている自治体データベース情報を活用して、得られた知見を報告する。

### B. 研究方法

**対象データ源とリスクアセスメント項目** 解析に使用するデータは、研究Ⅲと同様にして、三重県児童相談所のデータベースから抽出した。2014 年から、県独自に作成したリスクアセスメント項目への該当情報の電子記録が開始されている。リスクアセスメントは、通告による相談受付の単位で実施されている。実施のタイミングは、緊急受理会議時点と、一時保護判断時点があり、それぞれの段階でアセスメント項目への該当・非該当・不明が判断される。アセスメント項目は、主に緊急出動を検討する 6 項目と、一時保護の実施を検討する 15 項目、合計 21 項目から構成されている。

**データ源からのレコード抽出手続き** データ源からのレコード抽出は、児童相談所データベースに搭載されたリレーショナルデータベースマネジメントシステムから、リスクアセスメント関連テーブルのエクスポート機能を利用して実施した。抽出の対象期間は、2014 年 4 月 1 日から、研究実施時点最新月の 2021 年 12 月 09 日までとした。抽出対象は、受付の単位を 1 レコードとするリスクアセスメントテーブルとし、対象期間中に受理のあった児童虐待通告事例を抽出した。したがって、データ源からのレコード抽出に関する適格基準は、(1)指定期間内での受理であること、(2)児童虐待通告事例であることの 2 つとした。

**解析データの抽出手続き** データ源から抽出された指定期間内の虐待通告事例から、さらに解析に使用するデータ抽出を行なった。解析データの抽出に際して、虐待種別、児童の年齢に欠損がないことを適格基準とした。

**項目抽出とデータの前処理** 本研究では、一時保護の実施の有(1)無(0)を学習する正解ラベルとした。リスクアセスメント項目等の欠損値について、連続量の場合は中央値代入を実施し、離散量の場合は欠損ラベルを付与した上でOne-hot-encodingによるダミー変数に展開した。このとき、特に一時保護決定までの初動の段階では、得られる情報に限りがあることが多い。すなわち、リスクアセスメント項目への該当状況が「不明」という情報を活用するか、該当項目の累積等で「該当箇所のみに基づいて判断する」かについては、実践上の情報参照方式がことなる場合が想定される。よって、本解析では、「該当を1、それ以外を0」とする「該当情報のみを使用」する場合と、該当・非該当・不明のそれぞれをダミー変数化して不明情報を考慮する場合の2つの前処理を実施し、データセットを2つ準備した。リスクアセスメントテーブルから、最終的に解析に組み入れた項目を表4.1に示す。なお、リスクアセスメント情報の記録時点は、「一時保護判断時点」の情報を使用した。

表 4.1 組み入れ項目

変数	フィールド	処理
受付区分	新規・再通告・再受付	one-hot-encoding
主たる虐待者	実父・実母	one-hot-encoding
	実父以外の父・実母以外の母 その他	
主たる虐待種別	身体的虐待・性的虐待	one-hot-encoding
	心理的虐待・ネグレクト	
児童性別	男(0)・女(1)	正規化
児童年齢	0~18	
保護者性別	男(0)・女(1)	正規化
保護者年齢	0以上	
アセスメント項目1~21	該当・非該当・不明	label encoding (非該当と不明を0)
一時保護の実施	実施(0)・実施(1)	

**学習/検証データの分割** 学習データと検証データは、80%と20%の比率で無作為に分割して作成した。

**モデルと学習条件(XGBoost)** 一時保護の実施パターンを学習するモデルとして、研究Ⅱと同様に、XGBoostを使用した。XGBoostの主要な学習条件(ハイパーパラメータ)として、損失関数は対数損失(logloss)を使用し、学習率(eta)は0.03、条件分岐

の最大の深さは(max\_depth)4とした。ブースティングのラウンド数(num\_round)は、交差検証法(5-fold cross validation)により1回から8,000回の範囲で最適な値を決定した。その他の学習条件は、統計解析環境Rのxgboostパッケージにおける標準設定値を利用した(Chen, T. & Guestrin, C., 2016)。

**モデルと学習条件(バイズ多層ニューラルネットワーク)** 一時保護判断のパターンについて、予測時にそれが誤予測である可能性をあらかじめ評価するために、研究Ⅱおよび研究Ⅲと同様に、バイズ多層ニューラルネットワークモデルを適用した。本研究では、各中間層のノード数を全て36に固定した( $D_1 = D_2 = D_3 = 36$ )。パラメータ推定にかかるサンプリング条件は、iteration = 500, warmup = 250, chains = 4とした(total = 1,000 sample)。事前の誤予測を評価する際には、 $\theta_i$ の事後分布の幅の広さ(95%確信区間の上限と下限)を使用した。なお、バイズ多層ニューラルネットワークについては、XGBoostモデルとの比較を主眼とするため、不明を考慮した場合のデータセットにおいてのみ適用を行なった。

**性能評価** 予測性能を評価する指標として、研究Ⅱおよび研究Ⅲと同様に、PR曲線下面積(Area Under the Precision-Recall Curve: AUCPR)を採用した。本研究では、PR曲線下面積を積分によって求めた値を報告する。また、機械学習による予測と実際が整合していた割合を示す指標として、精度(Accuracy)も補足的に報告する。

#### (倫理面への配慮)

研究Ⅳは、人を直接の対象として協力や介入等を実施しないデータベース型研究である。当該研究に係る倫理審査については、下記の倫理委員会での審査を経て、承認を得ている。

【国立研究開発法人 産業技術総合研究所 人間工学実験委員会 事前申請】

実験課題名：児童相談所及び市区町村の児童相談データベースを活用したAI応用研究

判定結果：人間工学実験審査申請非該当(2020年3月9日)

【社会福祉法人恩賜財団母子愛育会 愛育研究所 研究倫理委員会】

受付番号：第4号

研究課題：児童相談所及び市区町村の児童相談データベースを活用したAI応用研究

判定結果：承認(承認番号第4号)

#### C. 研究結果

**組入データ** データ抽出の結果、18597件の事例デ

ータが解析に組み入れられた。学習データと検証データをそれぞれ 80%, 20%で無作為に分割した結果、学習データ 14877 件、検証データが 3720 件となった。それぞれに含まれたアウトカムの件数と該当率は、1903 件(12.8%)と 474 件(12.7%)であった。

**性能評価** XGBoost とベイズ多層ニューラルネットワークを用いた一時保護判断予測の結果、検証データを用いた際の AUC-PR が算出された。不明を考慮しないデータセットにおける XGBoost モデルによる予測では、AUC-PR が 0.66 となった。不明を考慮したデータセットでは、XGBoost モデルで AUC-PR = 0.70 の値が得られた。そして、不明を考慮したデータセットにおけるベイズ多層ニューラルネットワークの結果では、AUC-PR が 0.73 となった。それぞれの PR 曲線を図 4.1 に示す。

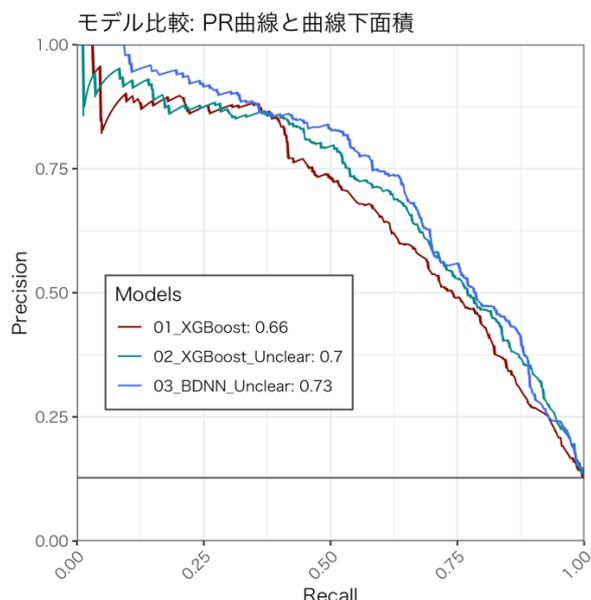


図 4.1 リスクアセスメント項目への該当状況が「不明」であることを考慮した場合と、「該当」情報のみを使用した場合の各種モデルの予測性能(PR 曲線)

続いて、もっとも AUC-PR の値が高かったベイズ多層ニューラルネットワークモデルにおける、検証データに対する予測を実施した際の予測スコア(出力層のパラメータ  $\theta_i$  に関する EAP)の分布を図 4.2 に示す。一時保護が実際に実施されなかった事例が 0 付近に多く分布し、実際に一時保護が実施された事例が 1 付近に多く分布しており、予測による識別性が高い様相が観察された。

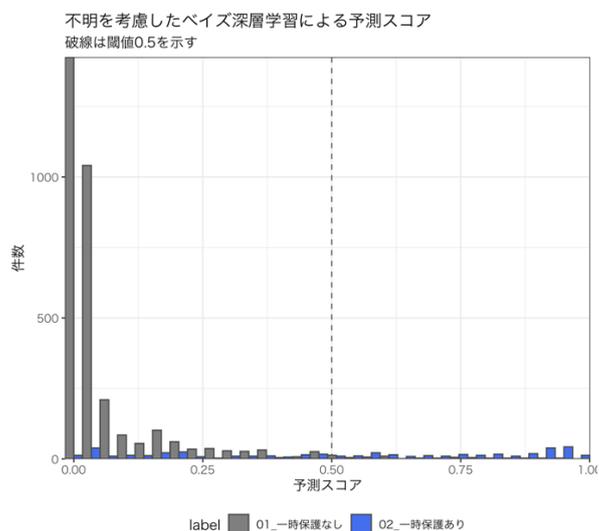
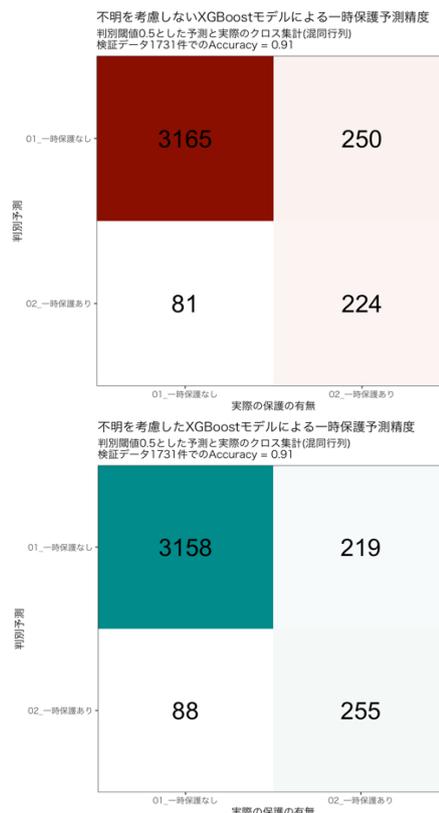


図 4.2 検証データを用いた出力層のパラメータ  $\theta_i$  の事後平均の正解ラベル別ヒストグラム

検証データを用いた予測の実施時に出力されたスコアの値に対して、閾値を 0.5 とする判別を実施した。その結果、いずれの解析においても、Accuracy は 0.91 となった。結果の混同行列を図 4.3 に示す。



不明を考慮したベイズ深層学習による一時保護予測精度  
判別閾値0.5とした予測と実際のクロス集計(混同行列)  
検証データ1731件でのAccuracy = 0.91

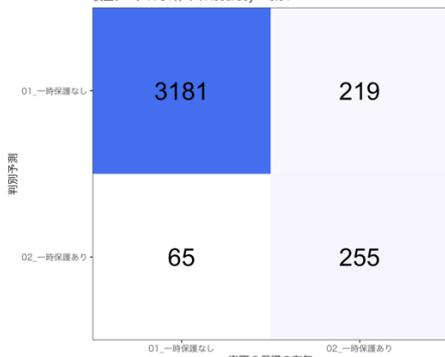
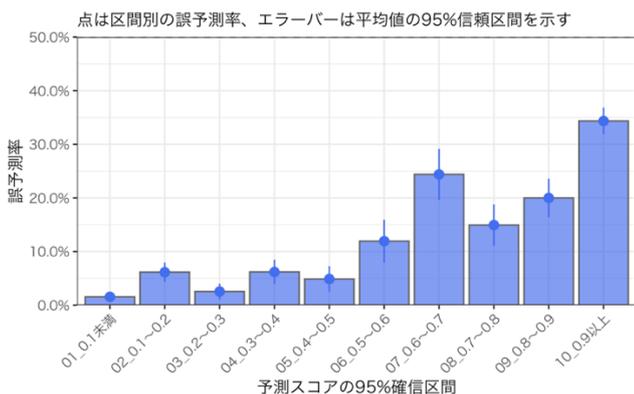


図 4.3 一時保護判断予測に関する 3つの解析結果の混同行列  
(検証データを利用, 閾値は 0.5 としている)

また、ベイズ多層ニューラルネットワークを用いた一時保護予測について、出力層のパラメータ  $\theta_i$  の 95%確信区間の幅の広さと、誤予測との関係を集計した。その結果、事後分布の幅の広さが大きいほど、誤予測率が高くなる傾向が観察された(図 4.4)。



#### D. 考察

本解析により、リスクアセスメントデータの該当・非該当・不明のパターンから、過去の一時保護対応履歴のパターンを学習し、新規事例に対して予測を出力することが可能であることが示された。比較的精度も高く、過去の判断の 91%程度が再現できることを示唆する結果が得られた。また、ベイズ多層ニューラルネットワーク等の技術を活用することで、誤予測の可能性も事前に評価できる可能性が示された。

日々対応される事例一つ一つに対して一時保護の判断を下す中で、過去の判断水準が維持されているかどうかを把握・調整することは容易ではない。ましてや、それが「年間」といった時間スケールで担保されているか否かを常時意識することは極めて困難となる。通告件数の増大など、刻一刻と変化する状況に対応する中で、判断基準の維持を図るためには、過去のデータの参照や、解析結果のフィードバックを逐次的に受け取る必要があ

るだろう。当該場面において、本研究で示した機械学習等の予測技術を活用した過去の対応パターンの予測的提示は、過去のデータから導かれたパターンをより正確な形式で参照する有効な手段となりうる。機械学習による予測そのものが、一時保護の決定根拠になることは妥当ではないと考えられるが(予測結果の理由が不透明であることや、「過去に類似パターンを持つ事例で保護していたから」という解釈は、目の前の対応事例の特徴に応じた保護理由を説明していないため)、多面的な情報を加味した判断を講じるなかでの一つの参照資源として有効に活用できる可能性がある。

また、本解析で示された「過去の判断パターン」を学習した機械学習等のモデルを活用することで、「将来的に虐待対応件数が増大していった場合に、過去の一時保護判断水準を維持したとすると、どの程度の一時保護需要が見込まれるか」といった予測を得ることも可能となる。児童の安全を確保してゆく上で、必要な施設環境等の支援資源がどの程度必要かを予測し、環境改善を施してゆく施策を講じる上での重要な判断材料となるだろう(一時保護の需要予測は研究Ⅷにて検討を行う)。個別事例に対する判断の補助だけでなく、施策的な意思決定への活用も含め、本解析で得られた知見の活用可能性は高いと考えられる。

#### E. 結論

リスクアセスメントデータを利用し、過去の一時保護の判断パターンを学習した解析モデルの構成は、個別事例に対する一時保護判断の補助や、これまでの一時保護判断の水準を維持した場合に必要な将来的な一時保護の需要予測(施設定員等の過不足・超過等)を予測するなど、様々な場面への活用可能性がある。意思決定の一貫性を確保し、必要な環境資源を確保する施策を講じるにあたり、一時保護の予測モデルの活用は重要な役割を担うと考えられる。

#### F. 健康危険情報

特になし

#### G. 研究発表

1. 論文発表  
なし
2. 学会発表  
なし

#### H. 知的財産権の出願・登録状況 (予定を含む.)

1. 特許取得  
なし
2. 実用新案登録  
なし

### 3.その他 なし

#### 【引用文献】

厚生労働省(2020). 令和元年度 児童相談所での児童虐待相談対応件数 (速報値). Retrieved <https://www.mhlw.go.jp/content/000696156.pdf> (2020. 12. 01)

厚生労働省(2019). 平成 30 年度子ども・子育て支援推進調査研究事業 事業報告書(2019).「課題番号 20: 里親担当児童福祉司、一時保護所の児童指導員等及び市区町村要保護児童対策調整期間職員の勤務実態に関する横断的全国調査と位置都道府県に対する縦断的 IoT センサーを用いたタイム スタディ」(事業主体 国立研究開発法人産業技術総合研究所).

Chen, T. & Guestrin, C. (2016). "XGBoost: A Scalable Tree Boosting System", 22nd SIGKDD Conference on Knowledge Discovery and Data Mining. <https://arxiv.org/abs/1603.02754> (Retrieved 2021. 05. 09).

## 研究 V

### A. 研究目的

児童虐待相談対応事例の数は年々増加しており、その背景に、DV・面前暴力による通告の増大があることなどが報告されている(山本, 2017; 厚生労働省, 2020). 全国あるいは自治体等の単位で、的確な施策や取組を講じるためには、増加する事例の内訳に関する情報を把握する必要がある。しかし、例えば重篤な事例の増減傾向や、子どもや家庭が抱える課題等に関する特徴別の増減傾向などは、本邦では統計が取得されてきていない。事例に関するアセスメント情報が電子的に処理・参照可能な形式で蓄積されていないことに由来するものと考えられる。

リスクアセスメント情報の蓄積が実施されれば、どのような虐待がどの程度発生しているかについて、その増減傾向等を含めた現状と将来の把握が可能となる。このとき、蓄積するアセスメントデータは、全国で共通した項目が含まれていることが望ましい。例えば、自治体別で傾向を比較することなどを通じて、各所の状況やニーズの差異が把握できるとともに、それに応じた個別の施策展開などにつながる可能性があるためである。また、把握可能な情報は項目内容に応じて広がることから、蓄積されるアセスメントデータは、可能な限り広範あるいは多面的であると良いだろう。

本研究では、自治体に蓄積されたリスクアセスメントデータを活用し、児童相談所が対応する事例の経年的な変化を捉える。その対象として、「一

時保護の実施可能性」に関する指標を事例ごとに算出し、その水準別での通告件数の変化の把握を行う。すなわち、「一時保護が高い可能性で実施される、重篤性や安全確保の必要性が高い事例」が増加しているのか、「一時保護の実施可能性が低い、比較的安全性の担保される事例」が増加しているのかについて、その傾向を把握する。

### B. 研究方法

**対象データ源とリスクアセスメント項目** 解析に使用するデータは、研究Ⅲおよび研究Ⅳと同様に、三重県児童相談所のデータベースから抽出した。同県には児童相談所が 6 箇所あり(うち、一箇所は 2019 年度新設)、各児童相談所での対応情報が一つのデータベース内に記録される仕組みとなっている。そして、全ての児童相談所にて、2014 年より、県独自に作成したリスクアセスメント項目への該当情報の電子記録が開始されている。リスクアセスメントは、通告による相談受付の単位で実施されている。実施のタイミングは、緊急受理会議時点と、一時保護判断時点があり、それぞれの段階でアセスメント項目への該当・非該当・不明が判断される。アセスメント項目は、主に緊急出動を検討する 6 項目と、一時保護の実施を検討する 15 項目、合計 21 項目から構成されている。

**データ源からのレコード抽出手続き** データ源からのレコード抽出は、児童相談所データベースに搭載されたリレーショナルデータベースマネジメントシステムから、リスクアセスメント関連テーブルのエクスポート機能を利用して実施した。抽出の対象期間は、2014 年 4 月 1 日から、研究実施時点最新月の 2021 年 11 月 31 日までとした。抽出対象は、受付の単位を 1 レコードとするリスクアセスメントテーブルとし、対象期間中に受理のあった児童虐待通告事例を抽出した。したがって、データ源からのレコード抽出に関する適格基準は、(1)指定期間内での受理であること、(2)児童虐待通告事例であることの 2 つとした。

**解析データの抽出手続き** データ源から抽出された指定期間内の虐待通告事例から、さらに解析に使用するデータ抽出を行なった。このとき、本解析では「通告件数(リスクアセスメント登録件数)」を主たる対象とすることから、解析データの抽出に際して、その予測数値に直接的な影響が出ないよう、除外基準は設けなかった。

**項目抽出とデータの前処理** 本研究では、一時保護の実施の有(1)無(0)を対象とした分析を行い、一時保護の実施可能性に関する指標を算出する(後述)。リスクアセスメント項目等の欠損値については、連続量(児童と保護者の年齢)の場合は、学齢区分・

年代等のカテゴリ変数を作成し、欠損の場合は「欠損」のカテゴリを組み込んでダミー変数に展開した。離散量の場合は、欠損ラベルを付与した上で、各カテゴリをダミー変数に展開した。なお、解析では一般化線形モデルを使用することから、各項目に含まれる第一カテゴリを参照変数(reference)とし、解析データセットから除外した。リスクアセスメントテーブルから、最終的に解析に組み入れた項目を表 5.1 に示す。なお、リスクアセスメント情報の記録時点は、「一時保護判断時点」の情報を使用した。

表 5.1 組み入れ項目

変数	フィールド
受付区分	新規・再通告・再受付
主たる虐待者	実父・実母 実父以外の父・実母以外の母 その他
主たる虐待種別	身体的虐待・性的虐待 心理的虐待・ネグレクト
児童性別	男(0)・女(1)
児童年齢	0~18
保護者性別	男(0)・女(1)
保護者年齢	0以上
アセスメント項目1~21	該当・非該当・不明
一時保護の実施	実施(0)・実施(1)

**統計解析(傾向スコアの算出)** 観察研究の枠組みで介入効果の評価を実施する際に、傾向スコア(propensity score)と呼ばれる指標が用いられることがある(Vandenbroucke et al., 2007)。傾向スコアとは、介入群・比較群の割り当てに関する可能性を定量化する指標であり、当該数値を用いて類似事例のマッチングや重み付けを行い、介入による因果効果等を補正的に推定する際などに活用される。児童虐待対応の文脈に照らせば、傾向スコアを用いて、一時保護に割り当てられる可能性を事例ごとに評価し、一時保護による再発防止の効果を推定する際などに活用することができる。本研究では、この「一時保護に割り当てられる可能性の度合い」(傾向スコア)を事例ごとに定量化し( $p_i$ )、その数値区分別で通告件数の増減傾向を把握する。傾向スコアの算出には、ロジスティック回帰を使用した。

$$p_i = \text{logit}^{-1} \left( b + \sum_{v=1}^V w_v \cdot X_{i,v} \right), \quad 0 \leq p_i \leq 1 \quad (5.1)$$

このとき、 $X_{i,v}$ は児童の年齢等の基本情報と、リスクアセスメント項目への該当状況を含んだ*i*行*v*列の行列となる。ここでは、全ての項目を傾向スコアの算出に使用した。なお、傾向スコアの値は、0 から 1 までの範囲の実数となる。本研究では、層別マッチングの手続きにならない、傾向スコアの値を用いて 0.2 刻みで 5 つの区分を定め、個別の事例

がどの区分に該当するかについて、カテゴリ化を行なった。

**統計解析(通告件数の増減傾向)** 一時保護の割り当て水準に関する指標である傾向スコアの区分別で、通告件数の増減傾向を把握するため、月次別での通告件数の集計を実施した。また、未観測の将来の通告件数を予測するために、本研究では、季節性調整項を組み入れた一般化線形モデルによる解析を行なった。月次(*t*)での通告件数を $Y_t$ ( $Y = 1, 2, 3, \dots, T$ )とし、1月から12月までの通告受理月を示すダミー変数の行列を、1月を参照変数として除外して $X_{t,m}$ ( $m = 1, 2, \dots, M$ )と表記する。そして、負の二項分布を目的変数に仮定した回帰モデルを次のように定義した:

$$\mu_t = \exp \left( \beta_0 + \beta_1 \cdot t + \sum_{m=1}^M \beta_{m+1} \cdot X_{t,m} \right)$$

$$Y_t \sim \text{Negative Binomial}(\mu_t, \alpha) \quad (5.2)$$

ここで、 $\beta_0$ は切片、 $\beta_1$ は月あたりの(対数スケールでの)増減量を示す。 $\beta_{2,3,\dots,12}$ は月別の季節性増減量を示す。なお、負の二項分布のパラメータは、ここでは平均( $\mu_t$ )と過分散パラメータ(dispersion parameter)を引数に定義した。なお、本解析モデルでは周期性成分以外のイベント効果等は考慮していない。したがって、新型コロナウイルス感染症の流行に伴った通告件数の変化等については、解析的には考慮せず扱っている。

(倫理面への配慮)

研究Vは、人を直接の対象として協力や介入等を実施しないデータベース型研究である。当該研究に係る倫理審査については、下記の倫理委員会での審査を経て、承認を得ている。

【国立研究開発法人 産業技術総合研究所 人間工学実験委員会 事前申請】

実験課題名：児童相談所及び市区町村の児童相談データベースを活用したAI応用研究

判定結果：人間工学実験審査申請非該当(2020年3月9日)

【社会福祉法人恩賜財団母子愛育会 愛育研究所 研究倫理委員会】

受付番号：第4号

研究課題：児童相談所及び市区町村の児童相談データベースを活用したAI応用研究

判定結果：承認(承認番号第4号)

C. 研究結果

傾向スコア区分別での件数トレンドを求める前

段として、一時保護の有無別で通告件数の増減傾向を式 5.2 のモデルで求めた。その結果、当該通告への対応では一時保護の実施されなかった事例の実件数が相対的に急な増加傾向を示し、一時保護の実施された事例の実件数は緩やかに増加している傾向が観察された(図 5.1 左パネル)。モデルによる推定値( $\mu_t$ および $e^{\beta_0+\beta_1 t}$ )を用いて、通告単位での一時保護実施率の推移を求めた所、季節性の影響を受けつつ保護率が緩やかに減少する傾向が観察された(図 5.1 右パネル)。

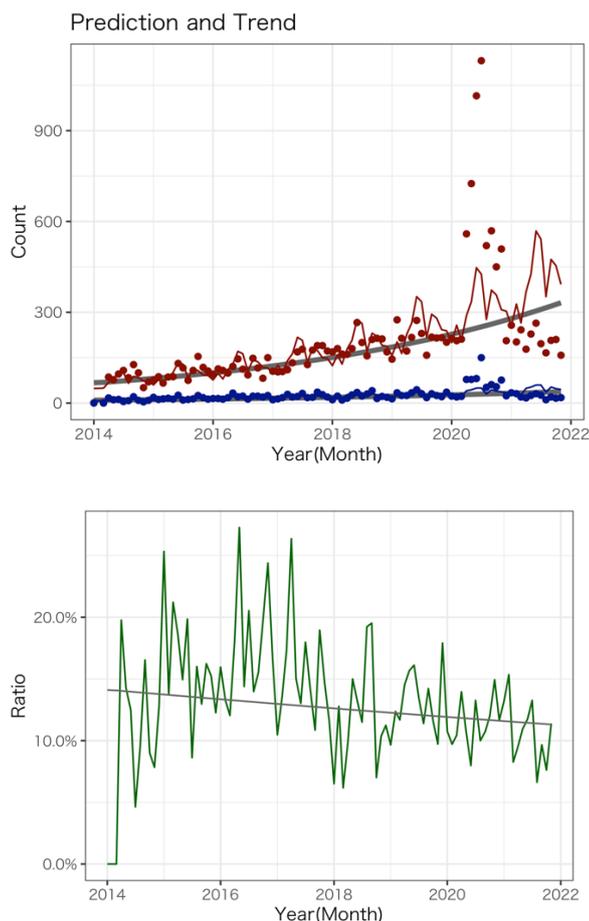


図 5.1 一時保護の実施別 通告件数と保護率の推定トレンド(月別)

点は実測値、実線は推定値を示す。灰色線は季節性調整項を除いたトレンドを示す。

上パネル: 赤色は一時保護の実施がない通告, 青色は一時保護の実施があった通告事例

右パネル: 緑線は、季節性調整項を含めたモデル推定値に基づく一時保護実施率の推移

続いて、ロジスティック回帰モデルで事例ごとの傾向スコアを求め、それを区分別に集計した基礎情報を表 5.1 に示す

表 5.1 傾向スコア区分別 該当件数と一時保護実施率

傾向スコア区分	該当数	一時保護数	一時保護率
$0.0 \leq p_i < 0.2$	15204	697	4.58%
$0.2 \leq p_i < 0.4$	1741	552	31.71%
$0.4 \leq p_i < 0.6$	652	315	41.31%
$0.6 \leq p_i < 0.8$	317	209	65.93%
$0.8 \leq p_i \leq 1.0$	683	604	88.43%

$0.0 \leq p_i < 0.2$	15204	697	4.58%
$0.2 \leq p_i < 0.4$	1741	552	31.71%
$0.4 \leq p_i < 0.6$	652	315	41.31%
$0.6 \leq p_i < 0.8$	317	209	65.93%
$0.8 \leq p_i \leq 1.0$	683	604	88.43%

傾向スコア区分別で層別化したデータセットに対して、式 5.2 に示したモデルを適用した結果を図 5.2 に示す。その結果、もっとも傾向スコアの小さい区分における通告件数の増加が顕著な様子が観察された。

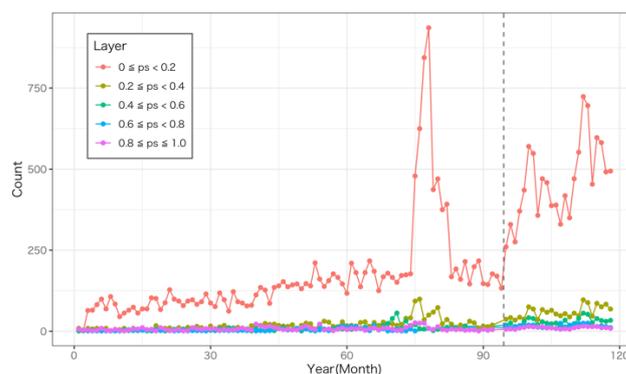


図 5.2 傾向スコア区分別 通告件数のトレンドと将来予測 (破線より左側がデータから学習したモデルのトレンドを示し、右側が予測値)

#### D. 考察

解析の結果、本研究で使用したデータベースを保有する児童相談所において、経年的な通告件数増加の背景に「一時保護実施率が相対的に最も低い群」の増加があることが示された。次いで傾向スコアの値が小さい群( $0.2 \leq p_i < 0.4$ )は将来的な増加が見込まれるものの、それ以上のスコアを持つ群には、他と比較して経年的な増加傾向がほとんど見受けられなかった。重篤性の低い事例など、在宅支援でも安全性が確保される事例への対応が当該児童相談所の件数上の中心となっており、その傾向が加速的に強くなる可能性がうかがわれた。また、新型コロナウイルス感染症流行期間におけるリスクアセスメント実施件数の増加など、将来予測のための解析では考慮されていない部分から、将来予測の値が大きく影響を受けた。当該観点については、様々な考慮のあり方が想定されるが、本研究の目的である「事例の水準別での将来的な需要が予測できることを示す」という前提に照らして、本研究内ではそれを扱わなかった。

比較的軽度な事例の通告増加が続いた場合、当該水準の事例対応に係る業務負担によって、間接

的に、重篤な事例への対応に負の影響が生じる可能性がある。すなわち、緊急性の高い事例や、児童の生命の危機が危惧されるような事例に対して、人的・時間的・環境的リソースを割くことができないといった事態の発生が危惧される。一時保護率の比較的少ない事例に対して、その一部に含まれる「一時保護による介入・調査が真に必要な事例」を的確に抽出するための仕組みや、それ以外的事例への対応に関する策を展開することで、業務負担の低減や的確なリソース配分の実現などが可能となるかもしれない。

#### E. 結論

リスクアセスメントデータを使用することで、端的に「どのような事例が近年増加しているか」等の情報を把握することが可能となる。本研究では、一例として、一時保護の実施率に関する指標を用いて、水準別での通告件数の経年変化の推定と将来予測を実施した。その結果、保護率の最も低い群の件数増加のみが顕著である様子が観察された。事例の内容等を反映するリスクアセスメントデータを使用することで、個別事例への対応に係る知見だけでなく、全国や自治体等で施策を検討する際の重要な参照情報を得ることもできる

#### F. 健康危険情報

特になし

#### G. 研究発表

##### 1. 論文発表

なし

##### 2. 学会発表

なし

#### H. 知的財産権の出願・登録状況（予定を含む。）

##### 1. 特許取得

なし

##### 2. 実用新案登録

なし

##### 3. その他

なし

#### 【引用文献】

厚生労働省(2020). 令和元年度 児童相談所での児童虐待相談対応件数（速報値）. Retrieved <https://www.mhlw.go.jp/content/000696156.pdf> (2020. 12. 01)

山本恒雄・高岡 昂太・久保 樹里・坂本次郎(2017). 平成 27 年度・28 年度 全国児童相談所長会 委託定例調査「虐待通告の実態調査（通告と児童相談所の対応についての実態調査）」.

Vandenbroucke, J. P., von Elm, E., Altman, D. G., Gøtzsche, P. C., Mulrow, C. D., Pocock, S. J., ... Egger, M. (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Medicine*, 4(10), e297. <https://doi.org/10.1371/journal.pmed.0040297>

## 研究VI

### A. 研究目的

蓄積されたリスクアセスメント情報は、基礎的な研究にも援用することができる。個別の事例に対する評価や組織単位の意思決定など、個別性あるいは事例性の高い問題だけでなく、場所や状況等に依存せず共有可能な知識の創出にも貢献する。長期的な視座から見て、一般化可能な知見を蓄積することの意義は大きい。「子ども虐待」に対する科学的理解の進展を含め、知識基盤の底上げに貢献する。

本研究では、リスクアセスメントデータを用いて、一般化可能性の高い知見の抽出を行う。主題として、「子どもからの保護の訴え」と「年齢」の関係を取り扱う。

子どもからの保護の訴えは、子ども虐待による死亡事例等を防ぐためのポイントとしても重視されている主要なリスク指標の一つである(厚生労働省, 2020)。保護を求める背景には相応の危機があると考えるのが自然であり、子どもの権利を保障する上でも、表明された意志の重要度は極めて高い。しかし、真に保護を必要とする状況であったとしても、子どもから必ず保護の訴えが得られるとは限らない。保護者を含めた家族等によって口止めされている場合や、「それが虐待である」ことを認識できない状況に置かれていること、あるいは子どもが複雑な葛藤を抱えていることなどによって、保護の訴えは困難となりうる。ここでは、保護の訴えが可能か否かを最も強く規定しうる子ども側の要因として、「年齢」を取り上げ、「保護の訴えが観測され始める年齢」の推定を行う。また、保護の訴えが可能となる年齢に関連しうる外的な要因として、虐待行為の内容、すなわち虐待種別を考慮に加えた検討を行う。

### B. 研究方法

**対象データ源とリスクアセスメント項目** 解析に使用するデータは、研究ⅢからⅤと同様に、三重県児童相談所のデータベースから抽出した。同県には児童相談所が6箇所あり(うち、一箇所は2019年度新設)、各児童相談所での対応情報が一つのデータベース内に記録される仕組みとなっている。そして、全ての児童相談所にて、2014年より、県独自に作成したリスクアセスメント項目への該当情

報の電子記録が開始されている。リスクアセスメントは、通告による相談受付の単位で実施されている。実施のタイミングは、緊急受理会議時点と、一時保護判断時点があり、それぞれの段階でアセスメント項目への該当・非該当・不明が判断される。アセスメント項目は、主に緊急出動を検討する 6 項目と、一時保護の実施を検討する 15 項目、合計 21 項目から構成されている。

**データ源からのレコード抽出手続き** データ源からのレコード抽出は、児童相談所データベースに搭載されたリレーショナルデータベースマネジメントシステムから、リスクアセスメント関連テーブルのエクスポート機能を利用して実施した。抽出の対象期間は、2014 年 4 月 1 日から、研究実施時点最新月の 2021 年 5 月 31 日までとした。抽出対象は、受付の単位を 1 レコードとするリスクアセスメントテーブルとし、対象期間中に受理のあった児童虐待通告事例を抽出した。したがって、データ源からのレコード抽出に関する適格基準は、(1)指定期間内での受理であること、(2)児童虐待通告事例であることの 2 つとした。

**解析データの抽出手続き** データ源から抽出された指定期間内の虐待通告事例から、さらに解析に使用するデータ抽出を行なった。なお、本研究では子どもの年齢と虐待種別を主たる変数として取り扱う。よって、当該項目のいずれかに欠損のあるレコードを除外の対象とした。また、性的虐待に関しては、蓄積されたデータサイズが小さい可能性が見込まれたため、あらかじめ解析の対象とせずに除外処理を行った。

**項目抽出とデータの前処理** 解析には、(1)子どもの年齢、(2)虐待種別、そして(3)「子どもによる保護の訴え」に関するリスクアセスメント項目を組み入れた。解析データは、虐待種別で層別化し、保護の訴えは該当(観測された場合)を 1、それ以外(明確に観測されていない場合)を 0 としてダミー変数化した。

**統計解析** 本研究では、「保護の訴えが観測され始める年齢」の推定に主眼を置いた統計解析を実施する。当該目的を達成するために、折れ線回帰モデルを基礎とした統計モデルを設定した。折れ線回帰モデルとは、一般に「ある分岐点より前のトレンドと、分岐点以降のトレンドが異なる」ことを仮定し、分岐点の位置と、その前後で異なる回帰モデルのパラメータを推定するモデルである。本研究では、ある年齢以前(分岐点 $c$ )では、保護の訴えが観測される確率( $p_i$ )が一定であることを仮定し(ノイズの考慮)、分岐点以降に観測確率が上昇することを仮定したモデルを構成した：

$$p_i = \frac{1}{1 + \exp(-\beta_0)}, \quad \text{if } X_i < c$$

$$p_i = \frac{1}{1 + \exp(-\beta_0 - \beta_1 X_i)}, \quad \text{if } c \leq X_i$$

$$Y_i \sim \text{Bernoulli}(p_i) \tag{6.1}$$

また、分岐点 $c$ が多峰性を有していた場合、保護の訴えが生じ始める年齢が段階的に存在する可能性が考えられる。そこで、推定時の提案分布に多峰性が認められた場合、一つ段階を増やしたモデルを適用することとした。分割点の数( $K$ 、ただし 2 以上の整数)が複数となる多段階の折れ線回帰モデルは、次式で定義した：

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \sum_{k=1}^K \beta_k c_k))},$$

$$c_k = X_i, \beta_1 = 0, c_k < c_{k+1}$$

$$Y_i \sim \text{Bernoulli}(p_i) \tag{6.2}$$

なお、モデル構成とパラメータの推定には、確率的プログラミング言語 Stan を使用した。各パラメータの事前分布は次の通りである。

$$c_{1,2,\dots,K} \sim \text{uniform}(1, 17)$$

$$\beta_{0,1,2,\dots,K} \sim \text{normal}(0, 10)$$

(倫理面への配慮)

研究VIは、人を直接の対象として協力や介入等を実施しないデータベース型研究である。当該研究に係る倫理審査については、下記の倫理委員会での審査を経て、承認を得ている。

【国立研究開発法人 産業技術総合研究所 人間工学実験委員会 事前申請】

実験課題名：児童相談所及び市区町村の児童相談データベースを活用した AI 応用研究

判定結果：人間工学実験審査申請非該当(2020 年 3 月 9 日)

【社会福祉法人恩賜財団母子愛育会 愛育研究所 研究倫理委員会】

受付番号：第 4 号

研究課題：児童相談所及び市区町村の児童相談データベースを活用した AI 応用研究

判定結果：承認(承認番号第 4 号)

### C. 研究結果

虐待種別で層別化したデータセットに対する解析の結果、身体的虐待とネグレクトを除き、心理的虐待では変化点が単峰に収束しなかった(1 点のみ変化点を仮定したモデルでは推定が成立しなかった)ため、変化点が 2 つのモデルを採用した。当

該条件における変化点の推定値を表 6.1 に示す。

表 6.1 子どもからの保護の訴えが増加し始める分岐点の推定値(虐待種別)

虐待種別	推定平均値 (EAP)	95%確信区間
身体的虐待	6.15(歳)	5.06-7.61
心理的虐待	7.26(歳)	5.25-8.72
	12.51(歳)	12.01-12.95
ネグレクト	11.72(歳)	7.04-13.97

変化点とその前後の推定トレンドを図 6.1 に示す。

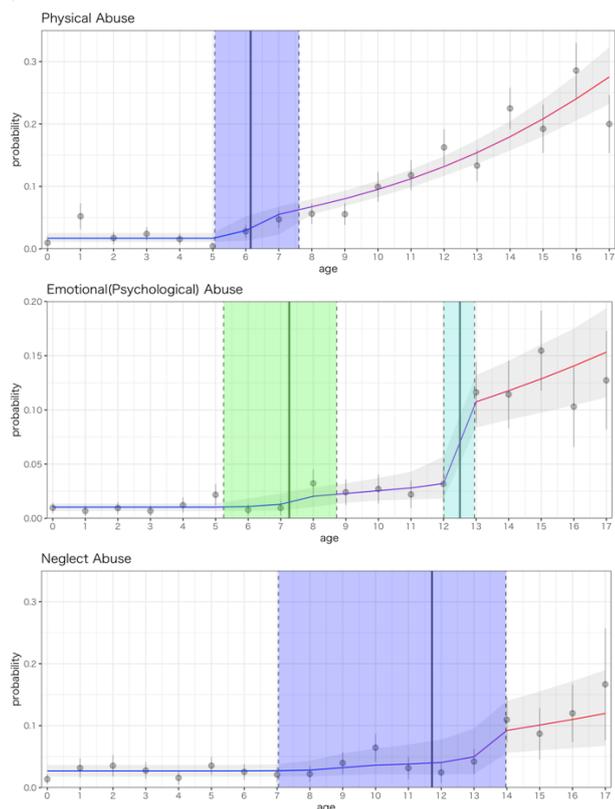


図 6.1 虐待種別 子どもによる保護の訴えの年齢別発生率とその推定値、変化点

#### D. 考察

子どもの保護の訴えに関する解析の結果、「保護の訴えが増加し始める年齢」は、虐待種別で異なることが示された。

身体的虐待については、小学校低学年ごろより増加し始める傾向が認められた。これは、「小学校

の段階では、中高生と同程度の水準で外部に SOS を発信することが難しく」、そして「小学校以前ではたとえ身体的虐待を受けていても周囲に保護を訴えることは基本的に困難」であるであると解釈することができる。

心理的虐待については、保護の訴えが増加し始める年齢が段階的であり、小学校低学年頃と中学校入学前後の二つに変化点があることが推定された。身体的虐待とは異なり、SOS の訴えが本格的に観測され始めるのは中学生頃以降となっていた。心理的暴力は、身体的暴力よりも SOS を訴えにくい性質を有している可能性がある。なお、心理的虐待では、保護の訴えが小学校低学年頃より一定の増加が認められる傾向も観察された。内容が深刻な場合や、一部の児童においては、小学校の段階から SOS が発信されはじめる可能性があると考えられる。ただし、全体的な傾向から、心理的虐待における保護の訴えは、中学生以上になるまでは基本的に難しいと捉えるのが自然だと言える。

ネグレクトでは、変化点の推定範囲が広く、明確な境界を求めることができなかった。ネグレクトを背景とする保護の訴えに関しては、年齢に伴う発達の变化とは異なる次元で発生している可能性も指摘される。ただし、全般的な傾向として、中学校以前の段階では、ネグレクトを背景とする保護の訴えが報告されにくい様相が観察された。

子どもからの保護の訴えは、虐待の内容や子どもの年齢によって制限を受けている可能性がある。すなわち、子どもが危機的な状況におかれていたとしても、SOS の訴えが得られない場合があると考えられる。当該知見は、慎重なリスク評価を実現する上で大きな示唆を有するものと考えられる。

リスクアセスメントデータの蓄積によって、「子ども虐待」に関する重要な知見が得られることの一例を示した。あらかじめ検証課題等を整理した上で蓄積データを設計することで、多様な知見の創出につながるものと考えられる。

#### E. 結論

リスクアセスメントデータを蓄積することにより、個別性や事例性の高い情報だけでなく、一般化可能性の高い学術的知見を得ることもできる。子ども虐待の重要なリスク評価指標の一つである「子どもによる保護の訴え」と「年齢」の関係を解析した結果、虐待種別で「訴えが報告され始める年齢」に違いがあることが示された。特に低年齢児童に対する慎重なリスク評価に際して、援用可能な知見が得られた。

#### F. 健康危険情報

特になし

#### G. 研究発表

1. 論文発表  
なし
2. 学会発表  
なし

#### H. 知的財産権の出願・登録状況（予定を含む。）

1. 特許取得  
なし
2. 実用新案登録  
なし
3. その他  
なし

### 【引用文献】

厚生労働省(2020). 子ども虐待による死亡事例等の検証結果について(第 16 次報告) 概要版, [https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000190801\\_00001.html](https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000190801_00001.html) (Retrieved 2021. 05. 27).

## 研究Ⅶ

### A. 研究目的

児童虐待事例の経年的な増加に対して、児童相談所の機能を維持してゆくためには、人的資源や環境資源等を含めた規模の拡大や、内部機能の効率化といった対処が迫られることとなる。この「効率化」について、「児童の最善の利益を優先する」という理念のもとでそれを捉えるならば、「子どもの最善の利益を確保した上で、対応に係る負担を最小化すること」を意味するものと考えられる。

児童虐待対応における「効率化」を実現する手段には様々なものが考えられる。例えば、事務的な手続きなどの対人援助に直接関わらない作業負担を ICT 等によって軽減したり、支援や介入の質を向上させる取組を講じたりすることなどがありうるだろう。

リスクアセスメントデータを活用するという視点から「効率化」を捉えた場合、例えば、「事例の終結までの期間を最小化する」といった課題設定ができる。このとき、終結時点の判断に用いられる主要な着眼点の一つに「終結しても再虐待が発生しない」という視点がある。虐待対応の理念と負担の軽減という両方を勘案すれば、「事例の特徴に応じた対応期間を設定することにより、再虐待の発生確率を減少させつつ、対応期間も短期化することは可能か?」という解析疑問が成り立つ。

本研究では、ベイズ深層学習と生存時間分析の手法を組み合わせた解析モデルを構築・活用し、事例の特徴に応じた対応期間別の再発確率を推定する。そして、再発確率と対応期間のトレードオフな関係を数量的に捉え、「現状の対応期間と再発

率の、両方を同時に改善可能か」について検討を行う。

### B. 研究方法

**対象データ源とリスクアセスメント項目** 解析に使用するデータは、三重県児童相談所のデータベースから抽出した。同県には児童相談所が 6 箇所あり(うち、一箇所は 2019 年度新設)、各児童相談所での対応情報が一つのデータベース内に記録される仕組みとなっている。そして、全ての児童相談所にて、2014 年より、県独自に作成したリスクアセスメント項目への該当情報の電子記録が開始されている。リスクアセスメントは、通告による相談受付の単位で実施されている。実施のタイミングは、緊急受理会議時点と、一時保護判断時点があり、それぞれの段階でアセスメント項目への該当・非該当・不明が判断される。アセスメント項目は、主に緊急出動を検討する 6 項目と、一時保護の実施を検討する 15 項目、合計 21 項目から構成されている。

**データ源からのレコード抽出手続き** データ源からのレコード抽出は、児童相談所データベースに搭載されたリレーショナルデータベースマネジメントシステムから、リスクアセスメント関連テーブルのエクスポート機能を利用して実施した。抽出の対象期間は、2014 年 4 月 1 日から、研究実施時点最新月の 2019 年 1 月 31 日までとした。抽出対象は、受付の単位を 1 レコードとするリスクアセスメントテーブルとし、対象期間中に受理のあった児童虐待通告事例を抽出した。したがって、データ源からのレコード抽出に関する適格基準は、(1)指定期間内での受理であること、(2)児童虐待通告事例であることの 2 つとした。

**解析データの抽出手続き** データ源から抽出された指定期間内の虐待通告事例から、さらに解析に使用するデータ抽出を行なった。解析データの抽出に際して、虐待種別、児童の年齢に欠損がないことを適格基準とした。また、研究Ⅲとは異なり、再発までの期間自体も推定対象とすることから、再発のラベル後に特定の観察期間等を設けず、したがって当該条件によるレコードの抽出や除外等は実施しなかった。ただし、「現状の対応期間」を扱うことから、エクスポート時点で未終結(対応中)にあるレコードは除外対象とした。

**項目抽出とデータの前処理** 本研究では、再発の定義を「通告の受付日から起算して、翌日以降かつ組入データの中でもっとも新しい日付までの間に、再度の通告記録が認められたレコード(当該事例が終結していたかは問わない)」とした。すなわち、同一の児童 ID がデータセットの期間内に反復出現し

たレコードを再発と定義し、再発ラベルを有(1)無(0)で付与した。このとき、「再発までの日数」を合わせて算出した。再発までの日数は、再発が観測された日から通告受付日までの日数の差分で定義された。具体的には、通告日の翌日に再度の通告が記録された場合、再発までの日数は1日となる。欠損への対処については、連続量の場合は中央値代入を実施し、離散量の場合は欠損ラベルを付与した上でOne-hot-encodingによるダミー変数に展開した。リスクアセスメントテーブルから、最終的に解析に組み入れた項目を表7.1に示す。なお、リスクアセスメント情報の記録時点は、「一時保護判断時点」の情報を使用した。

表 7.1 組み入れ項目

変数	フィールド	処理
受付区分	新規・再通告・再受付	one-hot-encoding
主たる虐待者	実父・実母	one-hot-encoding
	実父以外の父・実母以外の母 その他	
主たる虐待種別	身体的虐待・性的虐待	one-hot-encoding
	心理的虐待・ネグレクト	
児童性別	男(0)・女(1)	
児童年齢	0~18	正規化
保護者性別	男(0)・女(1)	
保護者年齢	0以上	正規化
アセスメント項目1~21	該当・非該当・不明	label encoding (非該当と不明を0)
一時保護の実施	実施(0)・実施(1)	

**学習/検証データの分割** 学習データと検証データは、80%と20%の比率で無作為に分割して作成した。学習データにおける再発のクラス不均衡に対しては、予測対象となるアウトカムが複数存在するため(再発の有無、再発までの日数)、対処等の手続きは実施しなかった。

### モデルと学習条件(対応期間別再発確率の予測)

本研究では、虐待事例への対応期間別での再発確率の予測を行う。解析には、バイズ多層ニューラルネットワークを基軸とし、生存時間分析の発想を組み入れたモデルの構築を実施した。本研究で使用したモデルを以下に定義する。

学習に使用する事例データセット $\mathbf{X}$ は、実数の $N \times D$ 行列である( $\mathbf{X}$ を構成する $N$ 個の行ベクトルは、それぞれ、 $D$ 個の項目からなる一つの事例データに対応する)。また、各事例における重篤な虐待の発生の有無を示す予測対象ラベル $\mathbf{Y}$ は、0 または 1 を成分とする $N$ 次の列ベクトルである。 $\mathbf{Y}$ の第 $i$ 成分 $Y_i$ の値が、 $\mathbf{X}$ の第 $i$ 行ベクトル $\mathbf{X}_i$ の表す事例データにおける再発の有(1)無(0)を表す。

本モデルで設定した中間層は3つある。各層に含まれるノード数をそれぞれ $D_1, D_2, D_3$ とする。また、 $D_0$ は入力層の次元 $D$ を表し、 $D_4$ は出力層の次元1を表すこととする。入力層から出力層までの重み付け係数を、それぞれ $\mathbf{W}^0, \mathbf{W}^1, \mathbf{W}^2, \mathbf{W}^3$ とする。各

$\mathbf{W}^k$ は、実数 $D_k \times D_{k+1}$ 行列である。各層における共通のバイアスを、実数 $b^0, b^1, b^2, b^3$ とする。

モデルの活性化関数 $f(z)$ は、次式で定義した:

$$f(z) = \frac{1}{1 + e^{-z}} \quad (7.1)$$

第一層から第三層までの中間層の各ノードで保持される値の組を、 $\mathbf{V}^1, \mathbf{V}^2, \mathbf{V}^3$ と表す。また、 $\mathbf{V}^0$ は入力層に与えられるデータセット $\mathbf{X}$ を表すこととし、 $\mathbf{V}^4$ は出力層で保持される値の組み $\theta$ を表すこととする。各 $\mathbf{V}^k$ は、実数の $N \times D_k$ 行列である。 $\theta$ は、 $N \times 1$ 行列、すなわち $N$ 次の列ベクトルで、各成分 $\theta_i$ に対して、 $0 < \theta_i < 1$ が成り立つ。各 $\mathbf{V}^k$ の $(i, j)$ 成分 $V_{i,j}^k$ の値は、事例データセット $\mathbf{X}$ の第 $i$ 行ベクトル $\mathbf{X}_i$ と、重み付け係数 $\mathbf{W}^k$ の第 $j$ 列ベクトル $\mathbf{W}_j^k$ を用いて、以下のように定義される:

$$V_{i,j}^{k+1} = f(b^k + \mathbf{V}_i^k \mathbf{W}_j^k), \quad k = 0, 1, 2, 3, \quad i = 1, 2, \dots, N, \quad j = 1, 2, \dots, D_{k+1} \quad (7.2)$$

ここまで、多層ニューラルネットワークによる事例ごとの推定再発確率( $\theta_i$ )が定義された。当該値は、「いつ再発するかを問わない、事例ごとの再発確率」の推定値である。これを対応経過期間別での値に変換するために、本研究では指数分布と、その累積密度関数を使用した。 $T$ は事例ごと経過期間( $T_i$ )を要素として持つ $N$ 次の列ベクトルである( $0 < T_i$ )。そして、 $\lambda$ は、その逆数が指数分布の期待値を示す $N$ 次の列ベクトルであり( $0 < \lambda_i$ )、対数リンクを介して実数の切片 $\alpha$ と、 $D_0$ 次の列ベクトルである回帰係数 $\beta$ から次のように定義した:

$$\log(1/\lambda_i) = \alpha + \sum_{j=1}^{D_0} \beta_j X_{i,j} \quad (7.3)$$

$$T_i \sim \text{Exponential}(1/\lambda_i)$$

そして、 $\mathbf{p}$ は経過期間に対応する $N$ 次の列ベクトルであり( $0 < p_i < 1$ )、要素 $p_i$ を「残存再発確率」と呼称し、次式で定義した:

$$p_i = \theta_i \cdot e^{-\lambda_i T_i} \quad (7.4)$$

モデルパラメータの推定にはハミルトニアンモンテカルロ法を使用した。各種添字を $i = 1, 2, \dots, N$ ,  $k = 0, 1, 2, 3$ ,  $l = 1, 2, \dots, D_k$ ,  $m = 1, 2, \dots, D_{k+1}$ としたとき、事前分布は、次のように仮定した:

$$Y_i \sim \text{Bernoulli}(p_i)$$

$$b^k \sim \text{Normal}(0,1)$$

$$W_{l,m}^k \sim \text{Normal}(0,1)$$

### (7.5)

なお、本研究では各中間層のノード数を全て 36 に固定した( $D_1 = D_2 = D_3 = 36$ ). パラメータ推定にかかるサンプリング条件は、iteration = 500, warmup = 250, chains = 4 とした(total = 1,000 sample).

事前の誤予測を評価する際には、 $p_i$ の事後分布の幅の広さ(95%確信区間の上限と下限)を使用した. 本研究での誤予測の定義は、「事例が再発すると予測した時に再発が未観測の場合、または、再発しないと予測した時に再発が観測された場合の、いずれかに該当する場合」とした.

**性能評価** 予測性能を評価する指標として、PR 曲線下面積(Area Under the Precision-Recall Curve: AUCPR)を採用した. AUCPR は、横軸に感度(Recall)、縦軸に適合率(Precision)を取り、機械学習が出力する予測スコアの判別閾値を変化させた時に現れる曲線(PR 曲線)で区切られた範囲下の面積を用いた指標である. 本研究では、PR 曲線下面積を積分によって求めた値を報告する. また、本研究のモデルでは、事例の再発までの期間を主要な予測対象としている. モデルによる予測と、実際の再発までの期間については、モデルから事後予測分布を生成して実際のデータとの対応関係を視覚的に確認する視覚的事後予測チェック(visual posterior predictive check)で整合性を確認した.

**再発確率と対応継続期間のシミュレーション** 対応継続期間と残存再発確率の関係から、許容される残存再発確率別での必要対応継続期間を算出した. すなわち、 $p_i$ が許容再発確率 $q(0 \leq q \leq 1)$ となるまでに必要となる(すなわち、 $p_i = q$ となるときの)対応継続日数 $\hat{T}_i$ を、式 7.4 より次式で求めた:

$$\begin{cases} \hat{T}_i = -\frac{1}{\lambda_i} \cdot \log\left(\frac{q}{\theta_i}\right) & \theta_i > q \\ \hat{T}_i = 0 & \theta_i \leq q \end{cases} \quad (7.6)$$

そして、 $\hat{T}_i$ を全事例で足し上げた値を評価指標として使用した(以下、総必要対応日数). すなわち、総必要対応日数  $S$  は次式となる:

$$S = \sum_{i=1}^N \hat{T}_i \quad (7.7)$$

(倫理面への配慮)

研究VIIは、人を直接の対象として協力や介入等を実施しないデータベース型研究である. 当該研究に係る倫理審査については、下記の倫理委員会

での審査を経て、承認を得ている.

【国立研究開発法人 産業技術総合研究所 人間工学実験委員会 事前申請】

実験課題名：児童相談所及び市区町村の児童相談データベースを活用した AI 応用研究

判定結果：人間工学実験審査申請非該当(2020年3月9日)

【社会福祉法人恩賜財団母子愛育会 愛育研究所 研究倫理委員会】

受付番号：第 4 号

研究課題：児童相談所及び市区町村の児童相談データベースを活用した AI 応用研究

判定結果：承認(承認番号第 4 号)

### C. 研究結果

**組入データ** データ抽出の結果、抽出期間時点で未終結であった 578 件を除く 7084 件の事例データが解析に組み入れられた. 学習データと検証データをそれぞれ 80%、20%で無作為に分割した結果、学習データ 5667 件、検証データが 1417 件となった. それぞれに含まれた再発事例の件数と該当率は、1432 件(25.3%)と 350 件(24.7%)であった. 再発事例における再発までの期間は、最小値が 1 日、中央値が 177 日、平均値が 276.3 日、最大値が 1614 日となった. なお、学習データにおける再発までの期間は、最小値が 1 日、中央値が 177 日、平均値が 278.1 日、最大値が 1550 日となった. また、検証データにおける再発までの期間は、

**再発予測の性能評価** 検証データを用いた再発予測に対する性能指標を求めた結果、AUC-PR は 0.754 の値が得られた(図 7.1).

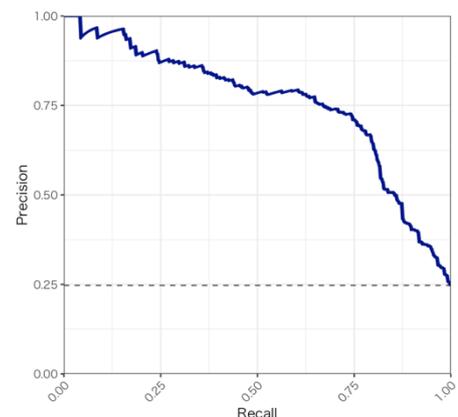


図 7.1 本研究のモデルによる再発予測時の AUC-PR

他の性能評価指標は、判別閾値を 0.5 としたときに、Accuracy = 0.865, Precision = 0.736, Recall = 0.709, Neg.Pred.Value = 0.906 となった.

**再発までの期間の推定** 検証データにおける、事例の再発確率( $\theta_i$ )、経過日数別の再発確率( $p_i$ )を図7.2に示す。推定の結果、個別の事例が有する推定再発確率 $\theta_i$ を切片として、経過時間別で残存再発確率 $p_i$ が減衰するトレンドが個別事例ごとに得られた。

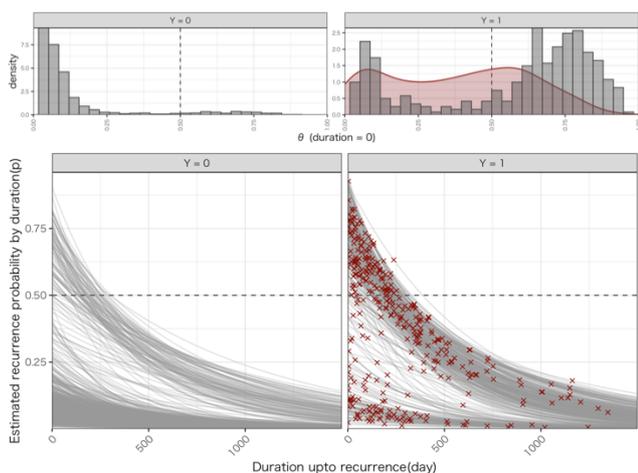


図 7.3 検証データにおける事例の推定再発確率と残存再発確率

付記されたヒストグラムは $\theta_i$ のEAPを示す。ヒストグラムに重ね書きされた赤色の密度曲線は、実際の再発時点における残存再発確率 $p_i$ のEAPを用いて描画している。トレンドを表記したパネルは、再発のなかった事例( $Y = 0$ )と再発のあった事例( $Y = 1$ )ごとに示した残存再発確率の推定トレンドである。赤色の×印は、再発事例における実際の再発が発生した経過日数時点で描画されている。

**視覚的事後予測チェック** 実際の再発事例における再発までの日数と、モデルで予測した再発までの日数分布を図7.4に示す。視覚的事後予測チェックの結果、実測値と予測値の分布が整合していることが確認された。

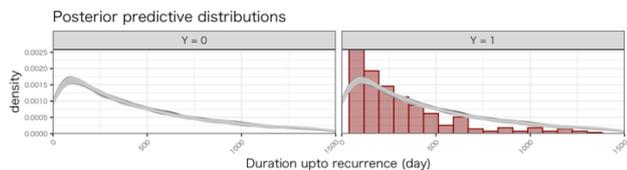


図 7.4 再発までの期間に関する事後予測チェック

**許容再発確率と総必要対応日数のシミュレーション** 許容再発確率 $q$ に0から1まで0.05刻みで値を与え、それぞれにおける総必要対応日数 $S$ を式7.5および式7.6により検証データで求めた。結果を図7.5に示す。なお、検証データにおける実際の再発確率と、終結までの期間(実際の対応日数)の総和を

図中に付記した(実際の対応日数の総和は、当該データ中に終結情報が含まれなかったことから、平均対応日数 = 335.9518 日と検証データの例数の積で求めている)。両結果から、再発率と総対応日数の両方が現状よりも小さくなる範囲があることが示された。

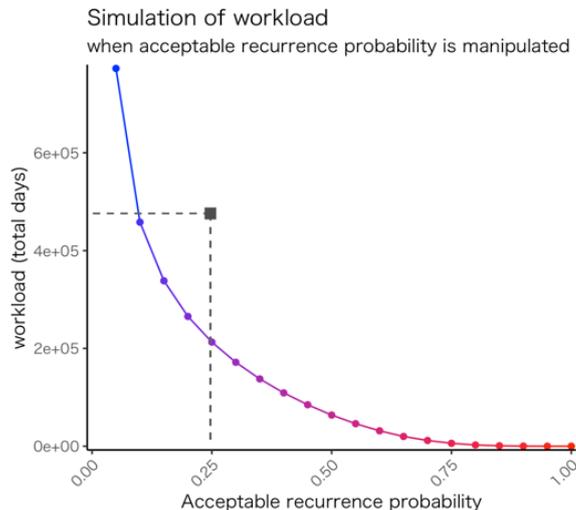


図 7.5 検証データにおける許容再発確率別での総必要対応日数

シミュレーション結果を折れ線グラフで示し、検証データにおける実際の再発率と総対応日数を黒点で示している。

#### D. 考察

解析の結果、再発確率と、再発までの期間が事例ごとに推定された。再発予測の精度も一定以上あり、再発までの期間に関する事後予測分布も実データと対応する結果となった。そして、許容再発確率を現状より低い値に設定し、それに至るまでの対応継続を事例ごとに実施した場合(ただし、その間は再発が発生しないという仮定を必要とする)、既存の対応日数と再発率よりも、短い期間で事例を終結でき、かつ再発率も低減できる可能性が示された。

リスクアセスメントデータは、一例として本研究に示したような形での援用につなげることができる。児童福祉の理念に応じた「効率化」を実現できる可能性を示した一例であると言えるだろう。しかし、本研究に示したような技術を実践的に活用する場合には、様々な課題が発生するものと考えられる。解析手法の内容や限界に関する適切な理解が求められることや、具体的な活用場面と活用する方法についても議論の余地がある。現状の業務や判断とどの程度の齟齬が生じるかなどについても、個別の検討が必要となるだろう。

#### E. 結論

蓄積されたリスクアセスメントデータを活用し、個別事例に対する再発の可能性を対応期間別で推定した。当該知見は、蓄積されたデータの高度な

活用方法の一例として参照することができる。シミュレーションの結果、既存の再発率や総対応日数の両方を同時に低減できる可能性が示された。データと解析技術は、児童福祉の理念に立脚した「効率化」を実現する手段の一つになると考えられる。

#### F. 健康危険情報

特になし

#### G. 研究発表

##### 1. 論文発表

なし

##### 2. 学会発表

Sakamoto, J., ○ Takaoka, K., Hojo, D., Hashimoto, E., & Furukawa, Y. (2019). Data-Driven Effective Case Management Method for Child Maltreatment with Bayesian Statistical Modeling and Machine Learning, International Society for the Prevention of Child Abuse & Neglect, Muscat, Oman.

#### H. 知的財産権の出願・登録状況（予定を含む。）

##### 1. 特許取得

なし

##### 2. 実用新案登録

なし

##### 3. その他

なし

## 研究Ⅷ

#### A. 研究目的

児童虐待対応業務に関する負担には、時間的負担や心理的負担等、様々な負担があるものと考えられる。中でも、広範な地域を管轄する都道府県下の児童相談所等では、家庭訪問等による移動コストが業務時間を圧迫することなどが想定されうる。児童虐待事例と地域の地理空間的情報を関連づけて捉えることは、こういった業務上の負担の把握や、地域の特徴に応じた予防的介入等の検討にも繋がるものと考えられる。

研究Ⅷでは、児童虐待事例の発生場所(住所地)と、児童相談所からの発生距離に関する分析を実施する。具体的には、自動車を移動手段とした場合に「何分程度で到達可能な地域か」を推定する到達圏解析を実施し、到達圏別での児童虐待事例の発生件数を定量化する。これにより、48時間以内の現認等で、仮に最寄りの児童相談所から家庭訪問等で現地まで自動車移動した場合に、累計でどの程度の移動時間が発生するか等について示唆を得る。

#### B. 研究方法

**データ源とレコード抽出** 解析に使用するデータは、三重県児童相談所のデータベースから抽出した。同県には児童相談所が6箇所あり(うち、一箇所は2019年度新設)、各児童相談所での対応情報が一つのデータベース内に記録される仕組みとなっている。データ源からのレコード抽出は、児童相談所データベースに搭載されたリレーショナルデータベースマネジメントシステムから、リスクアセスメント関連テーブルのエクスポート機能を利用して実施した。抽出の対象期間は、2014年4月1日から、2019年3月31日までとした(2019年度に1箇所の児童相談所が新設されたこと、解析に使用する道路・街路交通情勢の調査データが平成27年度のものであることから、対象を2018年度までに限定した)。抽出対象は、受付の単位を1レコードとするリスクアセスメントテーブルとし、対象期間中に受理のあった児童虐待通告事例を抽出した。したがって、データ源からのレコード抽出に関する適格基準は、(1)指定期間内での受理であること、(2)児童虐待通告事例であることの2つとした。

**解析データの抽出手続き** データ源から抽出された指定期間内の虐待通告事例から、さらに解析に使用するデータ抽出を行なった。解析データの抽出に際して、虐待種別と児童の住所地に欠損がないことを適格基準とした。

**到達圏解析** 平常時・混雑時の全国道路・街路交通情勢調査の結果を反映した、ESRI ジャパンデータコンテンツ ArcGIS Geo Suite(道路交通網情報を含む)を使用した。2018年度までに設置された県内5つの児童相談所の住所地を起点に、混雑時の自動車を想定して、15分区切りの最大120分以内で到達可能な範囲を県内にマッピングした。その後、児童虐待通告事例における児童の現住所地の空間情報と当該到達圏情報を連結した。このとき、事例の住所表記に係るクレンジング手続きにおいて、個人を特定しえない粒度に住所情報を抽象化した(町までを使用し、丁目や番地は除外した)。そして、各到達圏の通告件数を虐待種別ごとに集計することで、到達圏(地域)ごとの通告事例の様相を確認した。なお、これらの解析に際して、しよ島部の情報は道路ネットワークデータの対象外となるため、当該地域の情報および当該地域で発生した児童虐待通告の情報は除外されている。

(倫理面への配慮)

研究Ⅷは、人を直接の対象として協力や介入等を実施しないデータベース型研究である。当該研究に係る倫理審査については、下記の倫理委員会での審査を経て、承認を得ている。

【国立研究開発法人 産業技術総合研究所 人間工学実験委員会 事前申請】  
 実験課題名：児童相談所及び市区町村の児童相談データベースを活用した AI 応用研究  
 判定結果：人間工学実験審査申請非該当(2020年3月9日)

【社会福祉法人恩賜財団母子愛育会 愛育研究所 研究倫理委員会】  
 受付番号：第4号  
 研究課題：児童相談所及び市区町村の児童相談データベースを活用した AI 応用研究  
 判定結果：承認(承認番号第4号)

### C. 研究結果

データ抽出の結果、適格基準を満たした7903件の通告事例情報が解析の対象となった。

到達圏解析の結果、5つの児童相談所を起点とした場合に、県内の多くの地域が0分から60分以内に自動車で行き届くことが示された(図8.1)。その一方で、120分以内では到達できない地域も県内に存在することが示された。また、児童相談所ごとに管轄地域が異なることから、比較的短時間で移動可能な地域が多い児童相談所と、60分以上を要する管轄地域を広く含む児童相談所があることが示された。

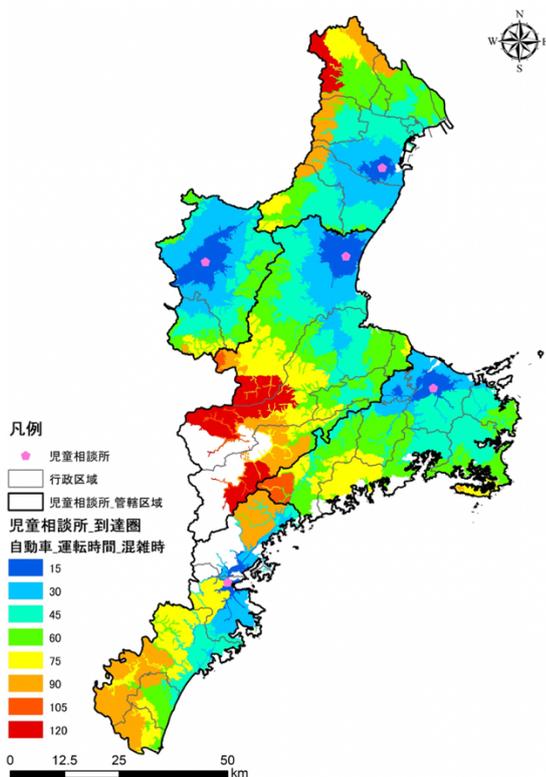


図8.1 2018年度までに設置のあった県内5つの児童相談所を起点とする混雑時の自動車による到達圏

※ 地図中の塗り分けは到達に係る所要時間を示すものであって、児童虐待事例の情報はここに含まれていない

次に、到達圏別で集計した児童虐待通告事例の該当件数を図8.2に示す。虐待種別での集計結果では、到達圏ごとにいずれの種別でも同様の傾向が観察され、児童相談所を出発地として30分から45分程度で現地訪問可能な範囲での通告事例が多く、60分からそれ以上の到達圏内での通告は相対的に件数が少ないことが示された。

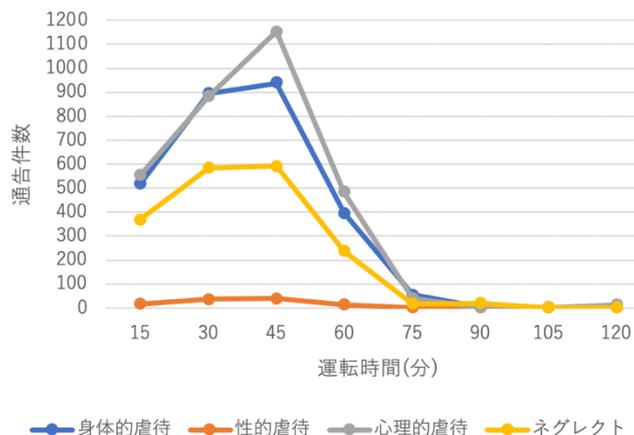


図8.2 到達時間区分別での児童虐待通告件数

※ 2014年度から2018年度までの、住所情報が正確に結合可能であった事例の情報のみを利用している(しよ島部の発生事例を除く)

### D. 考察

本研究では、児童相談所の設置住所から、通告のあった事例の住所地までの移動時間を到達圏解析で求め、各圏内での通告事例数を計上した。その結果、移動コストの高い地域を(面積的に)多く含む児童相談所と、比較的短時間で通告事例の住所地に訪問可能な地域を管轄する児童相談所があることが示された。当該結果は、ある都道府県を一例として扱ったものとなるが、全国の児童相談所で訪問等に係る所要時間等が様々に異なることを示唆するものであると考えられる。すなわち、「ある一件の通告」が持つ業務上の負担は、事例の内容等だけでなく「どのような地域や児童相談所の管轄圏での発生か、どの程度の移動時間が必要か」という観点からも異なるものと考えられる。通告件数や児童人口などの指標だけでなく、住所地までの距離(移動に係るコスト)等を考慮することで、各児童相談所での的確な人員配置の検討に繋がる可能性がある。

到達圏別での通告件数に関する結果から、現地訪問等の移動時間が累積的に膨大である可能性が推察された。なお、到達圏は「片道」に係る所要時間である。120分以上等の地域へと訪問した場合、往復で2時間程度以上の移動時間が発生することとなる。こういった業務時間の圧迫から、様々な意思決定や支援の内容への影響が発生しうる可能性も想定される。「状況の類似した事例に対して、異

なる到達圏で対応のあり方に違いがあるか」といった判断内容の均質性等に関する検討を加える上では、今後課題として、事例のリスクアセスメント等の情報を結合した上での詳細な解析が必要になる。

「全ての事例で、アセスメントの結果が必要な介入や支援に的確に結びつけられる」ことの実現を目指すにあたって、その阻害要因や促進要因を把握することは重要な課題である。その際、本研究で示した地理空間情報という視点は、移動コストや地域環境的な差異を考慮する上で、一つの重要な視点になるものと考えられる。

#### E. 結論

児童虐待通告事例の基礎情報と地理空間情報を結合して扱うことで、ある一つの通告に対する訪問等に係る移動コストが、通告のあった事例の住所や、管轄地域が異なる児童相談所によって異なる可能性が示された。移動に係る業務上の負担は、児童人口や通告件数などでは捉えきれない側面を反映するものであることから、的確な人員配置等を検討する上での重要な参考指標になる可能性が示唆された。また、地域ごとに児童虐待事例の性質がどのように異なるかを把握することは、地域を対象とした予防的介入を検討する上でも重要な情報となりうる。事例の詳細を含むアセスメント情報に対して、地理空間情報の視点を取り入れることで、データの活用内容と得られる知見の範囲に広がり期待できる。

#### F. 健康危険情報

特になし

#### G. 研究発表

##### 1. 論文発表

なし

##### 2. 学会発表

遠藤・坂本・北條・山本・高岡(2021). 児童相談所の管轄地域における児童虐待の地域的差異について, 2021年度日本地理学会春季学術大会, 2021年3月26日—3月28日.

#### H. 知的財産権の出願・登録状況 (予定を含む.)

##### 1. 特許取得

なし

##### 2. 実用新案登録

なし

##### 3. その他

なし

#### A. 研究目的

児童虐待に関する相談件数は年々増加の一途を辿る。並行して一時保護需要の増大が生じている。一時保護を担う受け皿(一時保護所定員・保護受託枠)には制限がある。現状の体制では、今後「本意ではない一時保護の未実施」や「本意ではない一時保護の短期化」の増大が懸念される。虐待防止という文脈からは、このような目先の対策は再虐待の可能性を増大させ、“かえって一時保護の需要件数を増大させる”という悪循環につながる恐れがある。今後の一時保護需要の増加に向けた、効果的な対策が必要になるだろう。

研究Ⅸでは、三重県内の児童相談所に蓄積された2019年度までのデータを使用し、以降の一時保護の需要予測を行い、その結果を報告する。「将来的な一時保護の必要件数と、一時保護委託がどの程度見込まれるのか」という課題に援用可能な知見を生み出すことを主眼としている。なお、本稿では「再発防止」の観点を取り入れ、長期的に効果の見込まれる形式(短期的な対策による悪循環や悪化の防止)での予測を行う。「再発防止に必要な一時保護日数」を解析し、その観点を含めた一時保護委託の将来需要を推定する。

解析は、大きく次の3つのステップで進める。第一に、再発防止への効果が期待される一時保護日数を推定する。これにより、「効果的な介入を実現するために必要な一時保護日数」が得られる。第二に、将来の一時保護件数を予測する。年々増加する一時保護委託件数から「将来一時保護委託が検討される件数」が得られる。上記2つの情報をもととして、「再発防止を考慮した日数を設定すると、将来的に同時に何件の一時保護委託が必要になるか」を推定する。

#### B. 研究方法

**解析の前提** 一時保護需要の将来推計を行う際には、いくつかの仮定を定める必要がある。本解析では、以下の条件のもとで推計を行った。

(1) 再発防止の有効性が認められるなかで、必要最小の一時保護日数を予測シミュレーションに採用する。

(2) 県内に設置された一時保護所における、現在の同時入所児童数水準を超えた一時保護需要は、全て一時保護委託での対応になると想定する(現状でもすでに生じている委託件数が増大傾向にあるため、保護所定員数を最大とせず、現状の一時保護所の入所体制を最大受け入れ体制であると仮定する)。

(3) 将来的な一時保護委託の需要予測は、「これまでの委託実施件数からの予測」+「再発を考慮した今後の保護需要推計から、一時保護所での受け入れをオーバーする件数」の2つを足し合わせ

## 研究Ⅸ

たものと想定する。

(4) 一時保護委託の需要を、件数ではなく「同時に委託する児童数」の形式で定量化する。

**データ源とレコード抽出** 「再発率を低減させる保護日数」の解析には、虐待対応に関するリスクアセスメントデータが必要になる。虐待対応に関するデータは、三重県に設置された5つの児童相談所が共通して利用するデータベースから取得した。具体的には、相談台帳、リスクアセスメント情報、一時保護台帳、措置台帳の4つのデータソースから虐待通告に該当するデータを抽出した。相談台帳には、児童相談所で対応のあった虐待通告を含む相談事例の情報が蓄積されている。性別や年齢等の基本情報は通告受理段階で入力され、一時保護の実施や終結に関する情報は、当該イベント発生段階で入力され、更新される。リスクアセスメント情報は、虐待通告時および現認時のリスクアセスメント情報と対応判断が、それぞれ通告受理、現認時、対応判断実施の段階で記録される。一時保護台帳および措置台帳には、児童虐待を根拠とした事例を含む、一時保護および措置に関する情報が記録されている。一時保護及び措置に関する情報は、その実施があった事例のみが記録され、実施および終結・解除時に情報が記録・更新される。

4つのデータソースに記録されたデータは、2018年9月6日までに三重県児童相談所のデータベースに蓄積された情報を抽出された。対象データは、2014年4月1日から2018年3月31日までに通告受理・対応記録が入力完了している虐待通告レコードとした。一時保護台帳に関するデータ(日別件数データ、件数予測に利用)は、2014年4月1日から2018年12月29日までの期間であった。なお、データ抽出時の段階で、氏名や住所等の個人情報項目は検索対象から除外された。以降の解析は、全てIDによる匿名化された状態で実施された。

**【統計手法1】** 状態空間モデルとロジスティック再発スコア調整を用いた一時保護期間別再発率の推定を実施した。三重県リスクアセスメントシートに含まれる情報を利用。1年以内の再発を予測するためのロジスティック回帰モデルを骨子とし、「リスク情報では説明が不可能な部分」を一時保護日数別で推定した。その際、「4日間の保護と5日間の保護の効果は似ている」「4日間の保護と6日間の保護はおおよそ似ている」といった、近い日数に再発への影響度の類似性を仮定する二階差分の状態空間モデルを適用することで、一時保護日数による再発率の変化を推定した:

$$\alpha_t = \text{Normal}(2\alpha_{t-1} + \alpha_{t-2}, \sigma)$$
$$p_{it} = \frac{1}{1 + \exp(\alpha_t + X\beta)}$$

$$\sigma \sim \text{Normal}(0,1)$$
$$Y_{it} \sim \text{Bernoulli}(p_{it})$$

ここで、 $Y_{it}$ は各事例における再発の有無を示し、再発確率は $p_{it}$ で示している。再発確率は、リスクデータ行列と各項目への重みベクトル $X\beta$ で予測され、シグモイド関数を介することで確率に変換される。この時リスクデータでは予測ができない部分を、 $\alpha_t$ で表記し、その生成メカニズムは一時保護日数 $t$ に制約をかけた状態空間モデルで推定される。なお、パラメータ $\sigma$ は状態空間モデルのシステムノイズと呼ばれる変動性の大きさを示すパラメータであり、一般的な仮定に合わせて事前分布にホワイトノイズ(平均0、標準偏差1の正規分布)を仮定した。パラメータ(一時保護日数別再発率)の推定には、ハミルトニアンモンテカルロ法を用いたベイズ推定法を利用した。

**【統計手法2】** 混合分布モデルを利用したパラメータ推定を実施した。一時保護の日数は、0日から2週間以内の割合が相対的に多く、長期事例の場合は60日を超過する。このような偏りのあるデータをモデル化するにあたっては、複数の確率分布を織り交ぜた混合分布を利用する必要がある。本稿では、短期間での一時保護解除の件数を負の二項分布で表現し、長期化する程度を離散一様分布で表現した。現状の一時保護日数を再現可能な要約モデルを作成することで、予測シミュレーションに利用する。保護日数のモデル化は確率的プログラミング言語STANを用いて実装し、負の二項分布の形状・尺度パラメータと、分布の切断点をベイズ推定法により求めた。

**【統計手法3】** 負の二項分布を用いた一時保護需要予測モデルを構成した。非負の整数データ(特に裾の伸びる分散の大きいデータ)を予測するために、一時保護件数の時系列変化を負の二項分布を用いた一般化加法モデルで学習・予測した。このとき、土日祝日の影響や月別の季節変動を考慮した:

$$Y_{it} \sim \text{NegativeBinomial2}(\beta_0 + \sum_{i=1}^{11} \beta_i^{\text{month}} + \sum_{w=1}^6 \beta_w^{\text{weekday}} + \beta_{\epsilon}, \alpha)$$

**【統計手法4】** 一時保護需要予測のモンテカルロシミュレーションを実施した。上記3つの統計手法を利用し、(1)将来の予測件数を生成、(2)必要な予測保護日数を生成、(3)保護定員の超過状況を計算、(4)一時保護の同時委託件数シミュレーションという4段階で、総合的なシミュレーションを実施。これにより、「再発率低減の観点を含め、今後どの程度の保護委託が生じうるか」について、2022年12月31日までの将来予測を行なった。

### (倫理面への配慮)

研究IXは、人を直接の対象として協力や介入等を実施しないデータベース型研究である。当該研究に係る倫理審査については、下記の倫理委員会での審査を経て、承認を得ている。

【国立研究開発法人 産業技術総合研究所 人間工学実験委員会 事前申請】

実験課題名：児童相談所及び市区町村の児童相談データベースを活用したAI応用研究

判定結果：人間工学実験審査申請非該当(2020年3月9日)

【社会福祉法人恩賜財団母子愛育会 愛育研究所研究倫理委員会】

受付番号：第4号

研究課題：児童相談所及び市区町村の児童相談データベースを活用したAI応用研究

判定結果：承認(承認番号第4号)

### C. 研究結果

状態空間モデルを用いた一時保護日数別再発率を比較した結果、「たとえ事例そのものが持つ再発リスクをリスクアセスメントデータから考慮した場合であっても、2日未満の一時保護は再発防止に相対的な効果が認められない」可能性が示された。以降の日数については、推定の結果明確な違いが認められなかった(図9.1)。

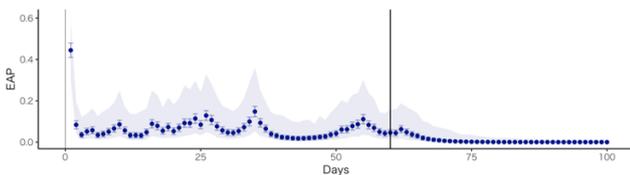


図9.1 リスクアセスメント項目で調整された保護日数別の再発確率の推定値

上段の解析では、「一時保護日数が1日の場合は再発率が相対的に高い」ことが明らかとなった。短期的な一時保護では、一時保護が持つ介入としての側面が弱まることによって、相対的な再発率の高さにつながっているものと考えられる。そこで、現状の一時保護日数の分布を基準に、「2日以上」になるような保護日数の調整を実施した。具体的には、現状の一時保護日数のパターンを、負の二項分布と離散一様分布の切断分布でモデル化した。そして、短期的な保護日数を表現する負の二項分布の期待値を2日分加算し、2日未満のシミュレーション値が得られた場合は再度のシミュレーションを実行することによって、「現状のパターンをもととした、2日以上」の保護日数分布」を生成した(図9.2)。

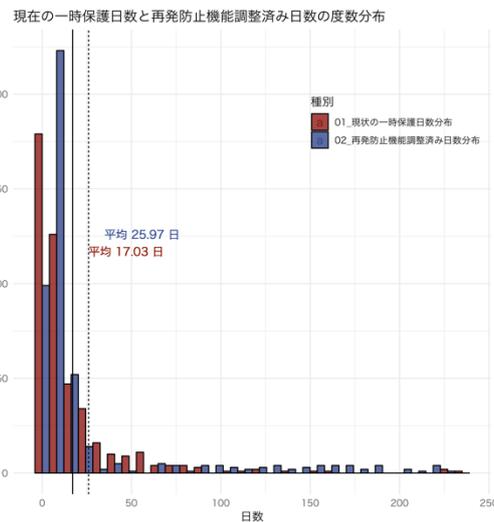


図9.2 一時保護日数の度数分布(赤)と、再発防止を想定した調整済み一時保護日数分布(シミュレーションによる一例)

再発防止を想定した一時保護日数のシミュレーションの結果、平均保護日数は従来の17.03日から25.97日に増大した。これはあくまで「平均値」であり、予測シミュレーションにおいてもその大半が14日未満の保護日数になることが想定されている。

続いて、一時保護の実施件数の予測結果を図9.3に示す。推定と将来予測の結果、季節性変動を含みながら、保護の実施件数は年々増大する傾向があると推測された。

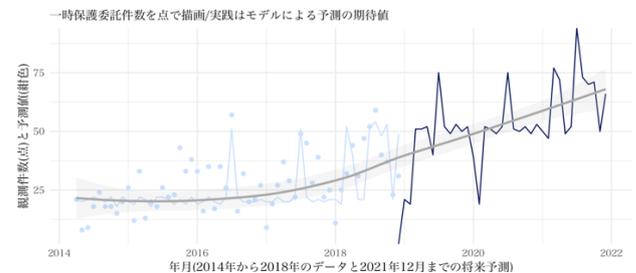


図9.3 季節性調整項を含んだ負の二項分布を用いた一般化線形モデルによる予測

この予測件数をもとに、現状の一時保護日数分布と、再発防止を考慮した保護日数の分布から、100回のモンテカルロシミュレーションにより同時保護人数を生成した結果、最大で31名の同時入所が予測対象期間内に発生する可能性が示された。当該予測結果と、県内児童相談所の受け入れ可能定員(非公開)の差分を「一時保護委託」として計上した。当該シミュレーションで得られた結果の一例(あるシミュレーションサンプル)を図9.4に示す。



図 9.4 モンテカルロシミュレーションを用いた一時保護委託の発生予測件数の一例

解析では、再発防止を考慮した一時保護日数を実現した場合に生じる一時保護所超過分を計上した場合、2019年では最大104件の同時委託需要、2020年には最大で97件の同時委託需要、2021年末までには最大で130件の同時保護委託需要が生じるという結果が得られた。なお、この数値は100回の乱数シミュレーションを行った結果の平均値を示すものであるため、およその参考値とされたい。

#### D. 考察

研究IXでは、リスクアセスメントデータの活用例として、再発防止への効果が期待される一時保護日数で一時保護を実施した場合に、どの程度の保護需要(同時保護人数)が将来的に発生するかが予測された。リスクアセスメントデータを活用することで児童相談所の主たる支援・介入の資源である一時保護機能に関して、将来的な需要を推測できる可能性が示された。

如何に優れたアセスメントツールや判断が実現されたとしても、必要な介入・支援を提供するための基盤に不足があれば、その効果は的確に得られない。こういった支援資源の必要性を定量的に示す上でも、リスクアセスメントデータが有効活用されうるものと考えられる。

#### E. 結論

リスクアセスメントデータは、事例のリスクを評価することだけでなく、様々な目的に活用することができる。介入効果が見込まれる対応方法を含めた一時保護の需要予測によって、必要となる支援資源に関する定量知見を得ることができた。当該手法を含めたデータの活用例は、各種行政施策等を策定する際にも援用可能であると考えられる。

#### F. 健康危険情報

特になし

#### G. 研究発表

#### 1. 論文発表

なし

#### 2. 学会発表

坂本・高岡・鈴木・山本(2019). 虐待対応に係る一時保護需要予測 — 再発防止を考慮した保護日数による同時保護件数のベイズ統計モデリング —, 日本子ども虐待防止学会第25回ひょうご大会.

#### H. 知的財産権の出願・登録状況 (予定を含む.)

##### 1. 特許取得

なし

##### 2. 実用新案登録

なし

##### 3. その他

なし

### 研究 X :

#### A. 研究目的

各自治体に赴き、3事例を提供してもらい、事例への評価結果と予測について現場感覚に合うかどうかについてヒアリングを行った。それにより、今後のAI実装を全国展開にさせる際のポイントを仮説生成した。

#### B. 研究方法

都市部、地方部を含む児童相談所と市区町村11自治体に連絡を行い、そのうち2自治体からはコロナ禍の対応のため、断られた。協力を得られた9自治体に赴き、産総研が作成したセーフティアセスメント(令和2年度厚労省子ども子育て調査研究事業成果物)に3事例を初期初動段階の情報を入力してもらった。そのデータを基に、産総研が開発したAIを用いて、一時保護の必要性、ケースの特徴を説明した。その結果、AIの予測結果と現場感覚が合うかどうかについてヒアリングを行った。

#### (倫理面への配慮)

産総研内の人間工学実験事前審査非該当、愛育研究所における倫理委員会の承認を得た。

#### C. 研究結果

9自治体の内訳は、都市部児相2箇所、地方部児相2箇所、都市部市区町村3箇所、地方部市区町村2箇所である。各自治体には、AIとは何かの説明、セーフティアセスメント自体の使い方説明、AIの予測する数値の読み解き方、どのように業務に参照したら良いかの説明を事前に丁寧に行った。各事例のフィードバックデータは個人情報を含むため、詳細は開示できない。概ね各自治体からはいただいた内容は、AIの予測結果と現場感覚がおおよそ一致するという評価であった。概要をまとめる

と、

- ・AIの予測と現場感覚が一致する点：
  - 保護の必要性、ケースの特徴については納得できる
  - マネジメントについては研究代表者の読み解き方の解釈も入っているが、それを含めて納得できる。
- ・AIの予測と現場感覚が一致しない点：
  - 初期初動の段階のため、不明な情報が多いケースの場合や、客観的かつ標準的なアセスメントではなく、ケースバイケースで判断するマネジメントが優先されたケースの場合は、予測結果と現場感覚が一致しないことがあった
  - 特に軽度ケースについては、セーフティアセスメントが目的とする一時保護の要否判断を満たさず、かつセーフティアセスメントでは対象としていない家庭のストレスやニーズまで現場は考慮しているため、リスクに応じて判断が一致しないことがあった。

#### D. 考察

高岡ら(2021)で提案された全国共通リスクアセスメント項目を現場支援者に実事例3例を評定してもらい、その評定を研究ⅢとⅣを基に開発したAIにデータを入力し、AIの予測結果を現場支援者へフィードバックした。その際の結果の納得感や使用感を現場支援者にインタビューした。その結果、現場支援者のAI出力に対する意見としては、「今まで現場ではデータを活用したことがなく、どのように数値を読み解くかといった知見がない」。そのため、「活用にはアセスメント結果だけでなく、どのように対応したらよいかのマネジメント側についても知見の提供や研修が不可欠」であることが明らかになった。

また、リスクアセスメント情報の継続的な蓄積は、子ども虐待対応領域全体の発展に寄与するものと考えられる。しかしながら、そのデータが適切に集めるためには全国的なデータの標準化が不可欠であり、AI等の予測技術を用いたリスクの評価だけでなく、必要な体制・支援資源の拡充根拠や、相談援助活動に携わる職員の業務負担軽減などを検討する根拠の創出にも活用するような伴走サービスが鍵になることが明らかになった。

#### E. 結論

リスクアセスメントデータは、児童虐待事例の詳細を標準的に捉えた情報源である。データ解析等での利活用を前提とした整然な形式で蓄積することで、透明性の高い客観的な根拠を創出することにつながり、領域全体の健全な発展に貢献しうる。定量的な視座から多くの知識を得ることで、子どもの最善の利益の追求がより確度高く実現されることにつながると考えられる。

#### F. 健康危険情報

特になし

#### G. 研究発表

##### 1. 論文発表

-Takaoka, K.,(2022). AI Implementation Science in Epidemiology and Observational Research: Pitfalls and Tips, Journal of Epidemiology.

##### 2. 学会発表

なし

#### H. 知的財産権の出願・登録状況（予定を含む。）

##### 1. 特許取得

なし

##### 2. 実用新案登録

なし

##### 3.その他

なし

#### 総合考察：政策への反映

AIを活用するためには、以下のような政策への反映が必要である。

①適切なリスクアセスメントのデータの入力支援施策

AI以前に全国基礎自治体や児童相談所に対するリスクアセスメントの研修や教育がこれまでされてこなかった。そのため、現状ではリスクアセスメントは現場では形骸化され、適切なデータが入力されていない。そのため、どのように使うかや入力するための基準の教育・研修が必要である。

②適切なデータを集めるためのシステム開発の  
リスクアセスメントの各データを集めるための項目の標準システムが必要である。現状では各自自治体が独自のリスクアセスメントを開発・運用し、システムも自治体の委託開発契約によりベンダーが

自由にシステムをカスタマイズしており、データとしての整合性がない。それらを統一するには、データベースの設計とAPI（Application Programming Interface）の実装で、個人情報を除くデータを収集できるような仕組みが可能となる。これを実装できればある程度データの標準化が進むと考えられる。

### ③AIを開発した後の現場への伴走支援

AIを開発しても、どのように使うかの現場への伴走支援がないと、結果として現場で使われなくなる可能性が高い。そのため、AIによる予測数値に対して、どのように現場で対応するかや各自治体の課題解決に対する伴走支援がAI実装と運用には重要となる。ここの伴走支援については各自治体

のニーズやデータ利活用のリテラシーレベルが異なるため、国全体で行うことは難しい。そのため、民間サービスなどと連携し丁寧に進める必要がある。④AIを活用する際に、エビデンスを参照した政策決定まで使うための政策設計

デジタル庁の子どもデータベースとの連携を含め、様々な領域との情報共有を前提に、現場に負担を掛けずにデータをA) 収集、B) 解析と実装、C) 政策に活かせる設計が必要である。A) データ収集には①で示した教育や研修と②で述べたAPIの実装が不可欠である。一方B) 解析と実装には③で記載した伴走支援、C) 政策への利活用にはAI用に集めたデータを政策にも利活用できる設計が重要となる。