

厚生労働科学研究費補助金（政策科学総合研究事業（政策科学推進研究事業））
（総括）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

研究代表者 高岡 昂太 産業技術総合研究所 主任研究員

研究要旨：

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、【研究 1】解析技術を用いたリスク評価と、【研究 2】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を行った。

研究 1：再発防止を考慮した同時一時保護
件数の推移

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。

本研究では、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から抽出された。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・対応記録が入力され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した。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基礎とし、「リスク情報では説明が不可能な部分」を一時保護日数別に推定した。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日間の保護はおよそ似ている」といった、近い日数に再発すること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、 α_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 σ 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(特に右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:

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

【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)。

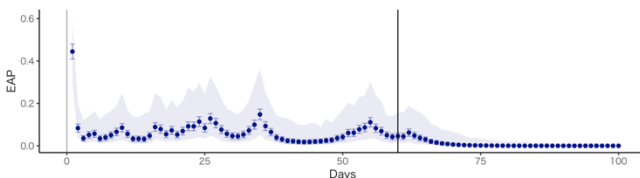


図9.1 リスクアセスメント項目で調整された保護日数別の再発確率の推定値

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)。

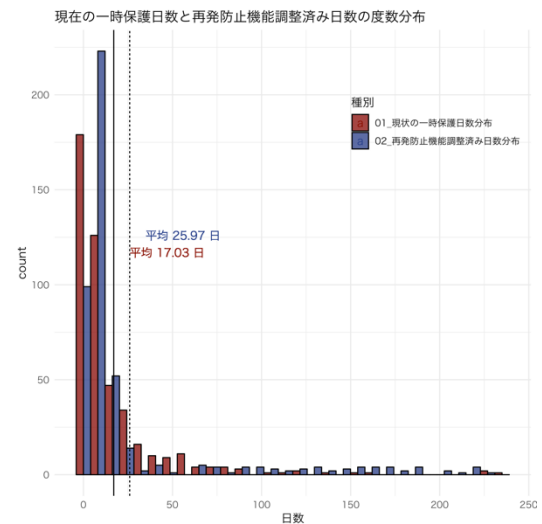


図9.2 一時保護日数の度数分布(赤)と、再発防止を想定した調整済み一時保護日数分布(シミュレーションによる一例)

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。

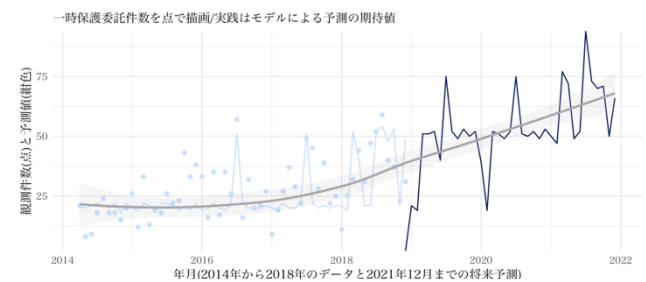


図9.3 季節性調整項を含んだ負の二項分布を用いた一般化線形モデルによる予測

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。

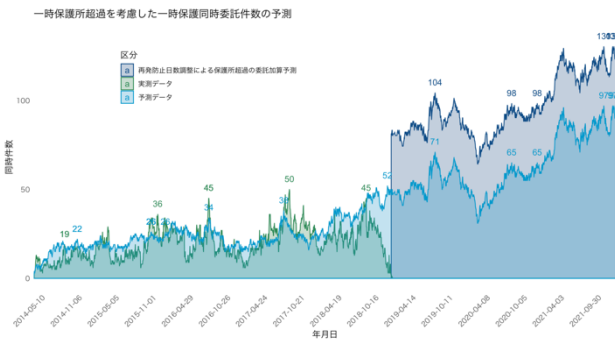


図 9.4 モンテカルロシミュレーションを用いた一時保護委託の発生予測件数の一例

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 2: リスクアセスメントツールの現場適用可能性

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ング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。

【考察】

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違うため、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。