

5.3.8 アセスメントツールの活用実態：アセスメントデータの集計・活用の実態

アセスメントデータを集計・分析・活用していると回答した組織の数は、アセスメントツールを利用している児童相談所119件のうちの12件(10.0%)であり、またアセスメントツールを利用している市区町村322件のうちの47件(9.6%)であった。

5.3.9 現在利用しているアセスメントツールの長所

現在利用しているアセスメントツールに関して、満足している点および使いやすい点などの長所について尋ねる質問の回答結果を表5.3.3にまとめた。アセスメントツールを利用している児童相談所119件から得られた主な意見は、「手間が掛からない」(38件)、「経験が浅くても使いやすい」(26件)、「認識を共有しやすい」(25件)、「判断に有用である」(21件)、「調査指針が明確」(14件)、「一枚のシートに収まっている」(11件)、「ケースを理解しやすい」(7件)、などであった。

アセスメントツールを利用している市区町村322件から得られた長所に関する主な意見には、「判断に有用である」(63件)、「認識を共有しやすい」(59件)、「手間が掛からない」(58件)、「調査指針が明確」(39件)、「経験が浅くても使いやすい」(37件)、「ケースを理解しやすい」(25件)、「客観的にアセスメントできる」(24件)、「一枚のシートに収まっている」(13件)、など児童相談所が挙げたのと同様の要素が挙げられた。さらに、市区町村に特徴的な意見として、「項目への該当を判断しやすい」(33件)、「客観的にアセスメントできる」(24件)が挙げられていた。

なお、長所に関するその他の特徴的な意見としては、「情報が少ない通告初期段階用のツールや送致用のツールなど、場面別のツールがある」(児童相談所2件)、「年齢別/虐待種類別/重篤度別ツールがある」(児童相談所1件、市区町村6件)、「ニーズやストレスを評価できる」(児童相談所2件、市区町村1件)、「ケースの変化を把握できる」(児童相談所1件、市区町村4件)、「養育者や児童と共有できる」(児童相談所1件、市区町村1件)、などが挙げられた。

表5.3.3 現在利用しているアセスメントツールの長所

	児童相談所(119件)	市区町村(322件)
手間が掛からない	38	58
経験が浅くても使いやすい	26	37
認識を共有しやすい	25	59
判断に有用である (含む:評価が数値化される) (含む:フローチャートで判断しやすい)	21 (3) (6)	63 (9) (7)
調査指針(確認すべき項目)が明確	14	39
一枚のシートに収まっている	11	13
ケースを理解しやすい	7	25
客観的にアセスメントできる	5	24
項目への該当を判断しやすい (含:手引きが分かりやすい)	4 (2)	33 (19)

5.3.10 現在利用しているアセスメントツールの短所及び課題

現在利用しているアセスメントツールに関して、課題を感じている点及び不満や不足などの短所について尋ねる質問に対する回答結果を表5.3.4にまとめた。

表5.3.4 現在利用しているアセスメントツールの短所及び課題

	児童相談所(119件)	市区町村(322件)
評価者によって評価が異なる	34	61
ケースの詳細が捨象される	11	16
フローチャートや数値化に基づく判断に不合理な点がある	8	4
手間が掛かる	6	33
項目への該当を判断しづらい	6	21
スキルを要する	3	4
年齢/虐待種別アセスメントができない	4	6
ネグレクトに関する項目が少ない	4	2
ストレングスが評価しづらい	3	1

アセスメントツールを利用している児童相談所119件から得られた意見には、「評価者によって評価が異なる」(34件)、「項目が限られているためケースの詳細が捨象される」(11件)、「フローチャートや数値化に基づく判断指標に不合理な点がある」(8件)、「手間が掛かる」(6件)、「項目への該当を判断しづらい」(6件)、「スキルを要する」(3件)、などが挙げられた。また、評価項目に関する意見として挙げられたのは、「年齢区分や虐待種別に相応しいアセスメントができない」(4件)、「ネグレクトに関する項目が少ない」(4件)、「ストレングスが評価しづらい」(3件)、などであった。さらに、児童相談所に特徴的な意見として、「どのような評価であっても一時保護を検討することになる」(3件)、「ツールの作成が意思決定の後となりアセスメントで利用できていないことがある」(2件)、が挙げられた。

アセスメントツールを利用している市区町村322件から得られた意見には、「評価者によって評価が異なる」(61件)、「手間が掛かる」(33件)、「項目への該当を判断しづらい」(21件)、「ケースの詳細が捨象される」(16件)、「ツールのみでは判断が難しいことがある」(6件)、「スキルを要する」(4件)、「フローチャートや数値化に基づく判断指標に不合理な点がある」(4件)、などが挙げられた。また、評価項目に関する意見としては、「年齢区分や虐待種別に相応しいアセスメントができない」(6件)、「ネグレクトに関する項目が少ない」(2件)、「ストレングスが評価しづらい」(1件)、が挙げられた。

上記で挙げた児童相談所及び市区町村からの主要な意見のうち、「評価者によって評価が異なる」並びに「項目への該当を判断しづらい」という課題に対しては、「状況例や程度の評価指標をより豊富に明示した記入要領やガイドラインを定めると良い」「評価の差を抑えるための研修がない/研修をする必要がある」という意見が10件挙げられていた。

なお、上記で挙げた主要な意見のうち、「ケースの詳細が捨象される」という意見27件は、「アセスメント項目外のリスク要因を反映しにくい」(3件)、「項目が少ないため精度に不安を感じる」(1件)、「項目が少なく実態に合わないことがある」(1件)、「項目が少ないため他に調査すべき状況がある場合にそれに目が向けられなくなる可能性を懸念する」(2件)、

「未知のリスクがあっても安全であるかのように見える」(1件)、「当てはまる項目がない時には判断しづらい」(11件)などから構成される。

さらに、特徴的な意見として以下のものが挙がっていた：

- ・利用方法に関する意見：「すべてのケースで利用できてはいない」(3件)、「形骸化している/アセスメントの補助具としての本来の使い方ができていない時がある/意思決定後の文書作成が目的となっている」(4件)、が挙がった。

- ・時間経過に関する意見：「2度目のアセスメントができていない時がある」(2件)、「どのタイミングで再アセスメントすべきか判断しづらい」(1件)、「どの時点の情報を記入するか迷う」(1件)、「各時点でリスクが異なり判断に迷う」(1件)、「どの時点でどの深さまで情報を取得する必要があるか悩む」(1件)、「ケースの状況の時間変化を追えない」(2件)、などが挙がった。

- ・評価や判断の数値化や自動化に関する意見：数値化や自動化のプロセスを含まないツールを利用している組織からは、「リスクが数値化されると対応を判断しやすい」(5件)、「自動的に判断を導出できると良い」(4件)、といった課題が挙がった。一方、数値化のプロセスを含むツールを利用している組織からは、「数値化されるが最終的な評価は職員の判断となる」(4件)という意見が課題や不満として挙がった。

- ・リスクの軽減要因に関する意見：「リスク要因と同時にリスク軽減要因がある場合の重篤度の評価に迷う」(1件)、「過去にあったリスク要因が現在は消失しているという状況に関する評価項目が足りない/そのような状況に該当する場合の緊急度の評価に迷う」(2件)、が挙がった。

- ・統計に関する意見：「列挙されているアセスメント項目の統計学的な妥当性が疑問である」(3件)、が挙がっていた。

- ・使いづらさに関する意見：「ケース記録や他のアセスメントツールなど複数の文書に同じ情報を記入する必要があり手間が重複する」(5件)、「アセスメント項目を追加できない」(2件)、「自由記述できない」(3件)、「記入欄が狭いなどの記入のしづらさがある」(2件)、「コンピュータでのExcelシートへの記入がしづらい」(1件)、「紙ベースなのでデータが蓄積できない」(1件)、などが挙がった。

5.4 結果：(B)アセスメントツールに対する一般的な意見

アセスメントツールを利用する上で特に重要視する要素及び最も重要視する要素は何かを尋ねる質問、並びに、アセスメントツールに対する意見を自由記述で尋ねる質問に対する回答結果を以下にまとめる。

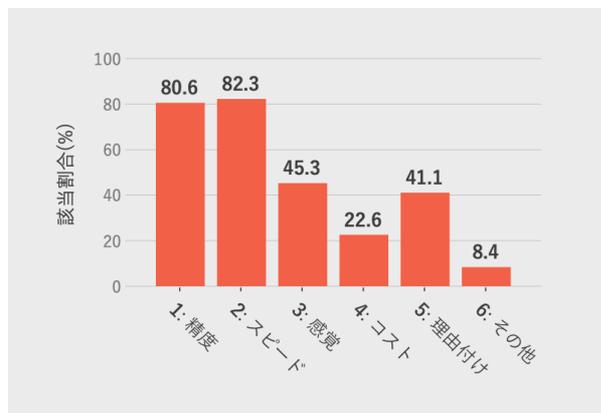
5.4.1 アセスメントツールを使用する上で特に/最も重要視する要素

アセスメントツールを使用する上で特に重要視する要素を尋ねる質問に対する回答結果を、児童相談所と市区町村の組織区分ごとに、アセスメントツールを利用している群と利用していない群に分けて集計した結果を図5.4.1に示す。なお、図に含まれていないアセスメントツール非利用の児童相談所1件が挙げたのは、「スピード」と「理由付け」であった。

アセスメントツールを利用している児童相談所と市区町村の回答結果は似た分布を示しており、精度とスピードを重要視する割合が約80%、次いで理由付けを重要視する割合が約4

0%であった。なお、児童相談所が「感覚(経験や直感に合うこと)」を重要視する割合は、アセスメントツールを利用している市区町村よりも13%ほど大きかった。

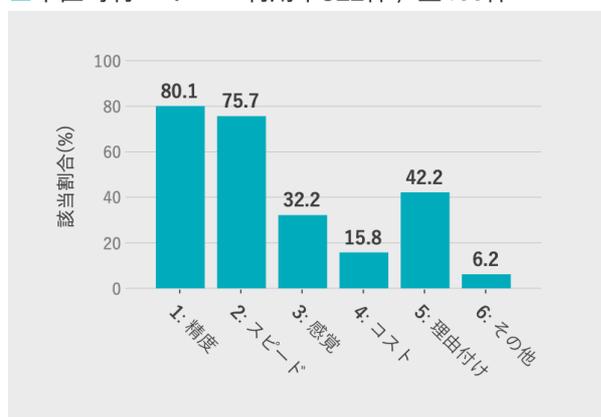
■児童相談所： ツール利用中119件 / 全120件



設問A2_11: リスクアセスメントを行う(リスクアセスメントシートを使用する)上で、特に重要視するものを全て選択して下さい。(複数選択)

- 1: 精度(児童虐待を予測・防止できるか)
- 2: スピード(迅速に判断できるか)
- 3: 感覚(経験や直感にどれだけ合うか)
- 4: コスト(業務量と導入による経済的部分)
- 5: 理由付け(「判断の際にリスクアセスメントシートを利用した」という実績)
- 6: その他

■市区町村： ツール利用中322件 / 全485件



■市区町村： ツール非利用の163件 / 全485件

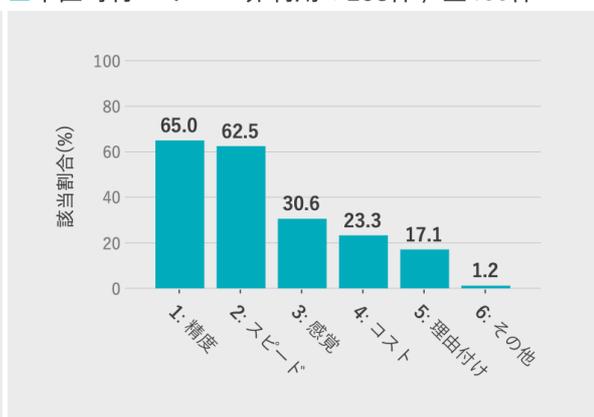


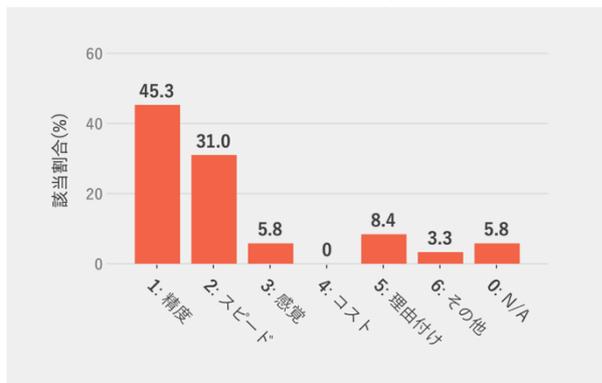
図5. 4. 1 アセスメントツールを使用する上で特に重要視する要素

一方、アセスメントツールを利用していない市区町村は、アセスメントツールを利用している児童相談所及び市区町村と比較して、理由付けを重要視する割合が顕著に低く、また精度とスピードを重要視する割合も幾らか低いという結果が得られた。

児童相談所から「その他」として挙げた主な要素は、「認識を共有しやすいこと」(3件)、「客観的に判断できること」(3件)、「評価者によって評価に差が出ないこと」(2件)、などであった。ツールを利用している市区町村から「その他」として挙げた主な要素は、「認識を共有しやすいこと」(7件)、「客観的に判断できること」(3件)、「評価者によって評価に差がでないこと」(2件)、「経験が浅くても使えること」(2件)、などであった。ツールを利用していない市区町村からは、「認識を共有しやすいこと」(1件)が挙げられた。

次に、アセスメントツールを使用する上で最も重要視する要素を尋ねる質問に対する回答結果を図5. 4. 2に示す。なお、図に含まれていないアセスメントツール非利用の児童相談所1件が挙げたのは「感覚」であった。

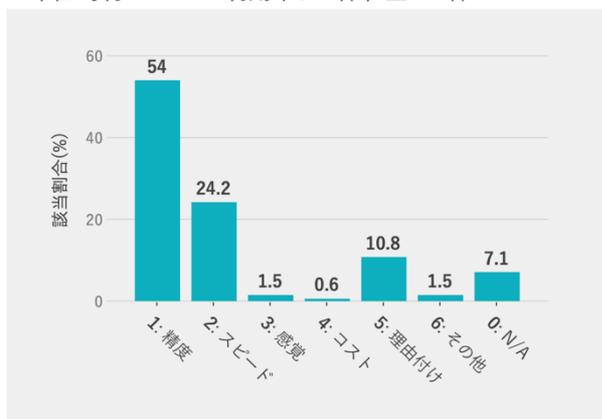
■ 児童相談所： ツール利用中119件 / 全120件



設問A2_11: リスクアセスメントを行う（リスクアセスメントシートを使用する）上で、最も重要視するもの一つ選択して下さい。（単一選択）

- 1: 精度（児童虐待を予測・防止できるか）
- 2: スピード（迅速に判断できるか）
- 3: 感覚（経験や直感にどれだけ合うか）
- 4: コスト（業務量と導入による経済的部分）
- 5: 理由付け（「判断の際にリスクアセスメントシートを利用した」という実績）
- 6: その他

■ 市区町村： ツール利用中322件 / 全485件



■ 市区町村： ツール非利用の163件 / 全485件

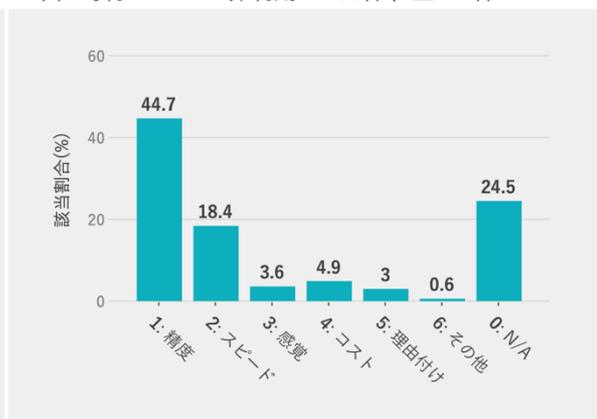


図5. 4. 2 アセスメントツールを使用する上で最も重要視する要素

ツールを利用している児童相談所119件のうち4件が、「その他」の要素を最も重要視していると回答した。その内訳は、「客観的に判断できること」（3件）、「認識を共有しやすいこと」（1件）であった。また、ツールを利用している市区町村322件のうち5件が、「その他」の要素を最も重要視していると回答した。その内訳は、「客観的に判断できること」（1件）、「認識を共有しやすいこと」（1件）、「組織内で判断を統一できること」（1件）、「調査や対応の進捗・時間変化を確認できること」（1件）、「改善を要する虐待要因が把握しやすいこと」（1件）であった。

児童相談所においても市区町村においても、アセスメントツールを使う上で最も重要視する要素は「精度」であると回答した割合が最も大きかった。アセスメントツールを利用している児童相談所と市区町村の回答結果は似た分布を示したが、精度を最も重要視する割合は市区町村のほうがやや大きく、一方でスピードを最も重要視する割合は児童相談所のほうがやや大きかった。また、図5. 4. 1と同様に、アセスメントツールを利用していない市区町村は、アセスメントツールを利用している児童相談所および市区町村と比較して、理由付けを最も重要視する割合がかなり小さかった。

5. 4. 2 アセスメントツールに対する自由記述意見

「アセスメントツールに関して感じること、思うこと、意見したいことがあれば自由に書いて下さい」という質問に対する回答のうち、主要な意見を以下にまとめる。この質問に対

する回答には、5.3.9項および5.3.10項で触れたアセスメントツールの主要な長所や短所及び課題に関する意見が多く並んだが、その他に見られた特徴的な意見を以下に挙げる：

- ・利用場面によって迅速さと精度の優先順位は異なる(児童相談所1件)：「緊急保護を検討する際には迅速さが重要であり、一時保護解除や措置解除を検討する際には精度が重要である。」

- ・利用場面ごとに確認すべきアセスメント項目を整理(児童相談所1件)：「通告受理時点での一時保護の要否判断、一時保護解除の判断、在宅指導終了の判断、市区町村への引き継ぎなど、それぞれの場面で確認すべき項目が整理されていると良い。」

- ・簡易迅速に利用できるチェック式と詳細を表現できる記述式の二段階のアセスメントツール(児童相談所1件、市区町村1件)：「迅速な対応が求められる場面で利用するツールと、迅速さよりも精度が優先される場面で利用するツールの二つがあって、状況に応じて使い分けられることができると良い。」

- ・初期対応時のアセスメントと初期対応後の支援段階のアセスメントそれぞれで利用できるツール(児童相談所2件)：「初期対応では迅速な対応につなげるためのツールが、支援段階ではニーズやストレングスも評価できるツールが求められる。」

- ・支援方針の判断に有用なツール(市区町村5件)：「リスクを評価した後に、どのような支援をどのような頻度で行うかなど、支援プランを検討できるツールがあると良い。」

- ・送致や役割分担に関する児童相談所の意見(2件)：「軽度虐待でも児童相談所が対応しなければならない実態がある/リスクの低いものは市区町村へ送致するという流れになると良い。」

- ・送致や役割分担に関する市区町村の意見(2件)：「重度項目に複数該当しないと児童相談所に送致を受けてもらえない/アセスメントツールを利用して一時保護などを考慮して児童相談所に送致してもあまり参考とされない場合が多いと感じる。」

5.5 結果：(C)厚労省共通ツールに対する意見

本調査では、厚労省共通ツールに対する意見として、当ツールを利用していない組織に対してはその理由を、当ツールを利用している組織に対してはその長所や短所及び改善のための意見を尋ねた。各質問に対して得られた回答のうち、主要な意見を整理した結果を表5.5.1に示す。

5.5.1 厚労省共通ツールの長所と短所及び改善に向けた意見

厚労省共通ツールの長所に関して、児童相談所から挙げた主要な意見は、「認識を共有しやすい」(6件)、「ケースを整理しやすい」(2件)、であった。また、市区町村から挙げた主要な意見は、「認識を共有しやすい」(16件)、「ケースを整理しやすい」(8件)、であり、さらに市区町村に特徴的な意見として、「客観的にアセスメントできる」(8件)、「調査指針が明確」(4件)、「経験が浅くても使いやすい」(2件)、が挙げた。

一方、厚労省共通ツールに関して、当ツールを利用している組織と利用していない組織両方から挙げた最も主要な短所及び課題は、項目が多く記述式である点に関連して、「手間が掛かる」(72件)であった。また、時間が掛かることに付随して、「認識を共有しづらい/

共有に時間が掛かる」という意見も2件挙げられた。市区町村に特徴的な意見としては、「職員に使いこなすスキルがない」(6件)が挙げられた。

また、当ツールの改善案として寄せられた主要な意見は、「項目を減らすなどして手間を減らす」(7件)、「判断基準をガイドラインで定めることや評価を数値化することによって評価基準や判断基準を明確化する」(5件)、であった。

表5.5.1 厚労省共通ツールに対する意見

厚労省共通ツールを利用しない主な理由

	児童相談所(120件)	アセスメントツールを利用している市区町村(322件)	アセスメントツールを利用していない市区町村(163件)
既存のツールを継続して使用している 又は県作成ツール等を利用している	90	193	
手間が掛かる	27	34	11
リスクの評価が曖昧	3	3	2
職員に使いこなすスキルがない		2	6
認識を共有するのに時間が掛かる	1	1	
当ツールについて周知されておらず知らない	3	8	20

厚労省共通ツールの活用しやすい主な点

	児童相談所	市区町村
認識を共有しやすい (含む:送致で活用しやすい)	6 (3)	16 (2)
ケースを整理しやすい	2	8
変化を把握できる	1	1
手間が掛からない	1	3
客観的にアセスメントできる		8
調査指針(確認すべき項目)が明確		4
経験が浅くても使いやすい		2

厚労省共通ツールの活用しづらい主な点

	児童相談所	市区町村
手間が掛かる	4	14
評価者によって評価が異なる	1	3
認識を共有しづらい	1	1

5.5.2 厚労省共通ツールの県改訂版の例

上記の課題を踏まえ、厚労省共通ツールをより活用しやすいように改訂したものを利用していると明示した組織が多数見られた。具体的には、本調査および前年度の調査(厚生労働省, 2020b)において、13箇所の都道府県における市区町村が「厚労省共通ツールを都道府県が改訂したもの」を利用していると回答した。これらの都道府県の多くは、表5.3.2において、アセスメントツールの利用率も共有率も高い都道府県であった。

改訂版の多くは、確認すべき項目を明示的に列挙して、その各項目に対して該当するか否か、あるいは確認できず不明であるか、また該当の程度といった情報を、記述式ではなく選択式で記入する方式を採用しているようである。例として、現在公開されている都道府県改訂版と厚労省共通ツールを図5.5.1に並べて示す。

5.5.3 厚労省共通ツールの利用実態に関するまとめ

5.3.5項で見た通り、どのアセスメントツールを利用しているかを尋ねる質問に対して「厚労省共通ツールを利用している」と回答した組織の割合は、アセスメントツールを利用

(11) 項目	(12) 状況例 ※別紙2「児童相談所と市町村の共通リスクアセスメントシート(例)」の記載上の留意点「リスク因子の主な指標例」を参照のこと	(13) 把握した状況及び様子	① 3/3		② 3/17		④【実施日 3/3】
			リスク		リスク		②【実施日3/17】
			あ	な	あ	な	受理会議等でのリスク
1	身体的な状況 (身体的虐待) ○頭部、腹部、胸部の殴打・蹴る等で生命の危機に係る受傷 ○受傷状況不明の骨折 ○首しめ・布団蒸し・鼻と口を塞ぐなど窒息につながる行為 ○乳幼児揺さぶられ症候群等の虐待による乳幼児頭部外傷疑い ○熱中症、低体温症を招くような環境下での放置 ○新旧差在した傷がある ○熱中症、低体温症を招くような環境下(車中の放置等)での放置 ○玄関やベランダに締め出し、子どもが求めても中に入れない等	① 通告受理日に学校に確認したところ、本児は変わりなく登校しており、見えるところに傷アザはなかった。 ③/3家庭訪問による現時点でも、本児に「親から叩いたり、暴力を振るわれることはない」との話を聴取。 ② 3/14 市子ども家庭課Wと共に家庭訪問時 ・母子双方から身体的な暴力がないことを確認。					① 学校での継続支援を依頼。 ② 1~2週間後に再度家庭訪問を行う。可能であれば、市にも同行を依頼。

項目	該当	該当なし	不明	現去は消失あり	状況例 (該当する項目があれば、チェックするか口で囲む。項目にないが特記すべき情報は余白に書き込む)				
					最重度	重度	中度	軽度	危惧
1				□	頭部外傷 乳児を投げる/踏みつけ 窒息の危険 その他生命危害行為	骨折 打撲 やけど 腹をける 顔面のひどい外傷 被害児が乳児	半年以内に2回以上のあざや傷 新旧の傷 顔面のあざ ける	傷が残らない程度の暴力 単発の小さくわずかなケガ	今傷はないが、発生する可能性が高い

図5.5.1 (上) 厚労省共通ツール, (下) 都道府県改訂版

している組織の数(児童相談所119件及び市区町村322件)を母数とした場合、児童相談所13.4%(16件)であり、市区町村31.3%(101件)であった。

ただし、この質問に「独自に開発したツール」またはその他のツールとして「都道府県で作成したツール」と回答した児童相談所は68件あった。また、その他のツールとして「都道府県で作成したツール」と回答した市区町村は103件あり、そのうちの23件は厚労省共通ツールを都道府県が改訂したものを利用している旨を明示していた。そして、表5.3.2に記載された共有率から分かるように、アセスメントツールを利用している市区町村のうちの76%が「児童相談所と共通利用している」と回答していた。また都道府県単位で見れば、アセスメントツールを利用している市区町村の3分の2以上の割合が「児童相談所と共通利用している」と回答した都道府県は、32箇所あった。

以上の状況を踏まえると、子どもの権利擁護の質を一層向上させるためには児童相談所と市区町村が円滑に認識を共有できることが重要であるとの厚労省共通ツールの目的は、多くの都道府県に理解されていると言えるだろう。

一方、厚労省共通ツールの様式や内容は、表5.5.1から分かるように、ケース情報を分かりやすく整理して表現する情報伝達ツールとしての側面が長所として評価されている一方で、アセスメントを補助するツールとしての側面に関しては、特に迅速な対応が求められる状況において活用しづらい点が短所として捉えられている。以上で述べた厚労省共通ツールの長所及び短所に関する主要な意見は、平成29年度に実施された児童相談所と市町村の共通アセスメントツール作成に関する調査研究事業(厚生労働省, 2018)においても複数のワーキンググループにより言及されていたものである。児童相談所及び市区町村がより実践的に活用できるアセスメントツールを構成するに当たっては、以上で述べた厚労省共通ツールの実情も踏まえて検討を行う必要がある。

5.6 考察

本章では、(A)アセスメントツールの活用実態、(B)アセスメントツールに対する一般的な意見、並びに(C)厚労省共通ツールに対する意見、を尋ねる調査の結果について見てきた。以下にその内容をまとめるとともに、その結果から得られた考察について述べる。

5.6.1 結果のまとめ

ほとんど全ての児童相談所がアセスメントツールを活用していた。特に一時保護の要否判断(96.6%)や家庭復帰の適否判断(60.5%)においてアセスメントツールを利用しており、厚労省一時保護シートを利用している割合も78.1%と高かった。比較的多くの児童相談所がアセスメントツールの運用方法にルールを設定しており(66.6%)、ルールに概ね沿った運用を行っている(72.2%)と回答した。

一方、市区町村は、児童相談所と比べると、アセスメントツールの利用率(66.6%)が低く、また運用方法にルールを設定している組織の割合(39.2%)や、概ねルールに沿って運用している組織の割合(47.5%)も比較的小さかった。

アセスメントツールの運用実態については都道府県ごとに特徴の差が見られた。市区町村におけるアセスメントツールの利用率や児童相談所とのアセスメントツールの共有率などには、都道府県による主導が大きく影響していると考えられる。

なお、アセスメントデータを集計・分析・活用している児童相談所及び市区町村の割合は、どちらも約10%であった。

アセスメントツールに求める機能的要素として児童相談所及び市区町村が共通して挙げたのは、「手間が掛からないこと」、「経験が浅くても使いやすいこと」、「認識を共有しやすいこと」、などであった。さらに市区町村は、児童相談所と比較して、「手引きが分かりやすいこと」、「確認項目への該当を判断しやすいこと」、「客観的にアセスメントできること」など、専門性の不足を補助する要素をより多く挙げた。

なお、アセスメントツールを利用する上で特に重要視する要素や最も重要視する要素を尋ねる質問に対して、アセスメントツールを利用している児童相談所及び市区町村は同様の回答傾向を示しており、精度とスピードを特に重要視していて、それに次いで理由付けも重要視していた。一方、アセスメントツールを利用していない市区町村は、ツールを利用している市区町村と比較して、理由付けを重要視する割合が顕著に低かった。

厚労省共通ツールに関しては、その内容や様式に対して、認識の共有に有用であるとの評価が挙げられた一方で、記述式であり項目が多いことから手間が掛かることを課題として挙げる組織が多かった。そのため、厚労省共通ツールを都道府県が改訂したものを利用していると回答した組織が多数見られた。

5.6.2 課題の整理

最後に、本章でこれまでに見てきた調査結果を踏まえながら、本研究の目的であるセーフティアセスメントツールの構成に関する課題を整理する。

第2章で述べた通り、本研究の目的は、前年度に作成した「児童虐待対応における市区町村・児童相談所で利用可能なセーフティアセスメントツール」を実践的に活用できるものとして構成し、それをセーフティアセスメントにおける「全国共通の物差し」として提案することである。

まず、初期対応段階での重篤事態の把握のために用いられるセーフティアセスメントツールが、より実践的に広く活用されるものであるためには、児童福祉に関わる上で最低限必要となる知見は研修などで習得することを前提として、(1)経験が浅くても利用できること、(2)手間を掛けずに迅速に利用できること、これら二つの要件を満たす必要がある。経験やスキルを要するアセスメントツールは、アセスメントツールを利用しない理由に関して述べた5.3.3項、ならびに厚労省共通ツールを利用しない理由に関して述べた5.5.1項で見た通り、それを「使いこなすスキルがない」組織、例えば人材確保やスキルの継承が容易でない小規模な市区町村などからは利用されづらいため、全国共通のツールとは成り難い。また、手間や時間の掛かるアセスメントツールが迅速な対応が求められる状況下では利用されづらいことも、先に挙げた5.5.1項を含め、本章において見た通りである。さらに、手間や時間の掛かるツールの活用は、5.3.10項で触れた通り、意思決定の後に事後的に作成されるという仕方、本来のアセスメントツールの利用法から逸脱し形骸化する恐れもある。したがって、セーフティアセスメントのための全国共通ツールは、例えば児童相談所の78.1%が利用している厚労省一時保護シートのように、経験が浅くても利用でき、なおかつ迅速に利用できるという要件を満たすものでなければならない。

さらに、本研究で構成するセーフティアセスメントツールは、児童相談所と市区町村で共通利用できること、すなわち、リスクの評価に関して、児童相談所と市区町村で共通の認識を得やすいツールであることが求められている。この要件を踏まえれば、職員間および組織間で評価に差が生じづらいツールであることを目指して本研究の目的であるアセスメントツールは構成されなければならない。5.3.9項で言及したように、市区町村は手引きが分かりやすいことやアセスメント項目の確認を行いやすいことをアセスメントツールを利用する際の長所として多く挙げていた。それを踏まえれば、評価に差の生じづらいアセスメントツールを構成するに当たっては、経験の浅い人でも各アセスメント項目を容易に確認することができるような手引きやガイドを構成し得るかについても同時に検討する必要があるだろう。

より実践的に活用できる全国共通のセーフティアセスメントツールを構成するには、以上で述べた要件を考慮する必要がある。5.5.3項で言及したような厚労省共通ツールの県改訂版もまた、以上の要件に応えるべく、記述式を該当/非該当/不明などの選択式に改め、該当するかを判断しやすい項目を調査項目として列挙する形式を採用するに至ったのだと推測される。

アセスメント項目が少ないことや記述式でなく選択式であることに対しては、5.3.10項で触れた通り、「項目が限られているためケースの特徴や詳細が捨象される」という課題も挙がっていた。しかし、本調査で得られたアセスメントツールの長所や短所に関する意見や厚労省共通ツールに対する意見を総合的に勘案すれば、本調査の結果は、特に虐待対応の初期段階におけるセーフティアセスメントの場面において、現場の職員の方々には概して、“一定の精度で”アセスメントを行えるという条件のもとで、経験が浅い人が利用しても評価に差が生じにくく、また可能な限り重篤性の高さを正確に判断できるツールを求めていると捉えることができるだろう。以上の要件を踏まえて、本研究において全国共通の物差しとして提案するセーフティアセスメントツールは構成される。

最後に、アセスメントデータの収集に関する課題について述べる。前段に述べた通り、全国共通のアセスメントツールは少なくとも“一定の精度で”アセスメントが行えるものでなければならないが、アセスメントツールの精度をより正確に評価し、またより良い精度のものに更新していくためには、より多くのアセスメントデータが蓄積されなければならない。しかし、5.3.8項で見たように、アセスメントデータを集計・分析・活用している児童相談所及び市区町村の割合は、アセスメントツールを利用している組織の約10%に過ぎない。また、5.3.1項で見た通り、市区町村のおよそ3分の1はアセスメントツールを利用していないのが現状である。すなわち、アセスメントデータを蓄積するのに十分な基盤が整っていない。そのため、本研究の提案するセーフティアセスメントツールを今後さらに良い精度のものに改善していくためには、業務現場のデジタル化の推進なども含め、より多くのアセスメントデータを効果的に蓄積するための基盤作りに向けた方策についても検討していく必要がある。そうした運用面に関する課題については、第10章及び第11章で改めて考察する。

第6章 全国調査結果(2)：アセスメントツールの評定者間一致性

【第6章の概要】

セーフティアセスメントツールを活用した児童虐待の一時保護判断において、評定者間で判断の不一致が生じることは望ましくない。どのような職員が保護判断を行ったとしても、「同一の事例に対して(概ね)同じ判断が下せる」ことがアセスメントツールとして目指すべき一つの理想となる。その際に重要となるのが、アセスメント項目に係る信頼性の検証である。

本章では、項目評定用仮想事例 (Web調査・パート2) を用いてセーフティアセスメント項目の信頼性について検討する。信頼性の評価指標としては、本事業では各アセスメント項目の「回答一致率」を使用する。同一の事例を評価対象とした際に、「回答者や組織によって評価が不安定な項目は、判断の不一致を生じさせやすい要因になる」という発想に由来している。

なお、本章では、回答一致率を用いた信頼性の基礎検証だけでなく、回答の不一致を生じさせている要因の分析も行う(潜在クラス分析、傾向スコアマッチング法を用いた2つの分析)。回答パターンの類型や、信頼性の低下に関わる要因が明らかになれば、その対策についても検討することができる。回答の不一致を防止する方法には、項目自体の表現等を工夫することや、評定者に対して「回答一致率を上げるために、どういった研修を行うことが必要か」といった観点からのアプローチなどが想定される。

6.1 目的：アセスメント候補項目の評定者間一致性の検討

6.1.1 一般的な信頼性の評価手続き

信頼性とは、調査項目の回答結果が一貫/安定しているかという観点から項目を評価する指標の一つである。信頼性が高い項目は、回答者(組織)内(間)で回答のブレが小さい項目として解釈することができる。「信頼性」は、様々な観点から評価が可能な多義的な用語だが、心理尺度等での信頼性を検証する場合は α (アルファ) 係数や κ (カッパ) 係数が用いられることが多い。また、信頼性の評価には時間的な安定性の観点(同じことを一定期間のうちに実施しても、回答が一致するか等)から検討されることもある。

6.1.2 子ども虐待対応に係るアセスメントツールの信頼性評価について

本研究のように児童虐待対応に関わるアセスメントツールの信頼性評価を行う場合、 α 係数もしくは κ 係数のような指標を用いる場合もある。しかしながら、児童虐待対応で用いられるアセスメント項目は様々な観点から子どもの安全を評価するものであり、上記の信頼性係数のみで評価することは困難である場合が多い。専門的な用語で言えば、複数の構成概念で作成された項目群を単純な因子構造に適した信頼性係数で評価することは難しいためである。

複数の構成概念で選定された項目群に対しては、因子ごとに信頼性係数を算出することも可能ではあるが、性格特性等を測定する尺度と異なり、児童虐待対応に係るアセスメント項目は対象とする事例の特徴に応じてその評価が異なることもしばしばあり(ex. 屋外に放置されていたというアセスメント項目に対し、夏季の沖縄と冬季の北海道の事例ではその評価の重みが異なることは言うまでもない)、信頼性係数が本来想定している評価基準にならないことがある。このような事例の特徴によって評価が異なることは最初に述べたアセスメント項目を評価する際の回答者(組織)のブレとは異なるものであり、本調査で検討する信頼性(ブレ)はあくまでも後者の回答組織間(内)の評価のブレであることを明記しておく。

上記を踏まえ、本研究におけるアセスメント項目の信頼性の評価には、回答組織間のブレを評価するために、信頼性係数のみによる評価だけでなく、単純集計による項目回答の一致率を項目ごとに算出することによって信頼性の評価を行った。

6.1.3 項目評定用仮想事例を用いた場面想定法の活用

前項で述べたように、児童虐待に係るアセスメントでは、事例の特徴に応じて評価が異なることがある。そして、回答者に由来して評価の不一致が生じることもある。本研究では、特に後者について、「同一事例に対して回答者(組織)間で安定した評価が可能か」について検討することが目的である。この目的を達成するために、本研究では、提示するアセスメント項目に対応した事例(以下、項目評定用仮想事例)で評価を行う場面想定法を採用する。これを用いてアセスメント項目の回答一致率を算出し、その信頼性について検討する。

6.2 方法

6.2.1 項目評定用仮想事例の作成手続きと有識者評価

場面想定法によるアセスメント項目評定に用いる項目評定用仮想事例は、計12事例作成した(研究デザインの詳細は、3.3.2項に記載している)。項目評定用仮想事例作成の手順は、まず、現場で児童虐待対応経験のある研究者が、本研究で検討するアセスメント項目を各仮想事例で評価可能となる(該当・非該当の判断がつけられる)ような12事例を作成した。その際、仮想事例の特徴(児童性別・児童年齢・虐待種別・きょうだいの有無等)は、全12事例を通して可能な限り偏りがないように配慮を施した。そののち、心理臨床経験のある研究者および臨床経験のない(項目の信頼性について適切な統計処理を行うことが可能な)研究者の数名が、各アセスメント項目との対応を想定している仮想事例において、該当または非該当の評価が可能かどうかを、実際に評価することで確認した。この確認のプロセスを通して、仮想事例の内容や各回答グループに提示する項目の調整を行った。最後に、これらの仮想事例を、検討委員会(児童虐待対応現場の職員・児童虐待防止に関する研究者で構成された組織)に参加している有識者で評価し、そこでの議論を基に修正を行い、本調査で使用した。実際に使用した項目評定用仮想事例については巻末に記載している。

有識者から提出された主要な意見の中には、各項目評定用仮想事例に関する指摘も多数含まれたが、3.3.2項にも記載していたように、本調査を実施する際の言葉遣いに関するものもあった。例えば、本調査で使用した事例は元々「仮想事例」という表記を用いて調査を実施

する予定であったが、実際に上記の手順で作成された事例を確認した際に、「今回用意した事例は、実際の児童虐待対応事例とはほど遠いと判断される可能性がある。「仮想事例」という表記を改める必要がある」という指摘があった。これを踏まえ、「項目評定用仮想事例」という表記に改めることを決定した。そして、回答組織に教示する文章で項目評定用に作成した仮想的な事例のため、実際の児童虐待対応事例と一部異なる可能性がある旨を記載した。

6.2.2 データ処理・集計方法

本調査データを3.3.2項の手順で収集した後、単純集計による回答一致率の算出および信頼性係数としてクリッペンドルフの α 係数(Krippendorff's alpha coefficient)を算出した。クリッペンドルフの α 係数は、内的一貫性を検討するクロンバックの α 係数とは異なる。この指標は、様々な評定者間一致性に関する信頼性係数(ex. コーヘンの κ 係数等)をより一般化したものであり、1(完全一致)から、-1(完全不一致)の数値を取る。信頼性がない(一致しない)場合は0の値を取り、比例・間隔・順序・名義尺度のいずれにも拡張可能な指標である。これらの算出した指標を基に、項目の信頼性を評価した後、回答者属性による回答傾向の検討を実施する。

6.2.3 回答者属性による回答傾向の比較

前項の基礎集計結果を踏まえ、回答者属性に応じて項目評定の不一致が説明可能かを検討する。本調査では、回答者の経験年数という要因に特に着目し、解析を実施した。回答者の経験年数に加えて、回答者の職位やこれまでの経歴なども考慮すべきであるだろうが、調査全体の回答負担および各回答組織における同名職位による仕事量の違い等を考慮し、今回の調査では、最も測定容易かつ結果に影響を及ぼしそうな回答者属性である経験年数にのみ着目した。

本研究では、まず、(1)潜在クラス分析を用いた、回答者の経験年数による回答パターンの分類を実施し、その結果を踏まえ、(2)回答者の経験年数が項目評定に影響を与えているかについて、傾向スコアマッチングを用いた因果推論を実施した。

6.2.3.1 潜在クラス分析

本研究では、各仮想事例に関して各組織から得られたアセスメント結果において、特に20個の必須アセスメント項目(巻末資料参照)に注目し、その20個の項目全体を通して現れる傾向を検出して考察することを目的として、潜在クラス分析(Lazarsfeldら, 1968)を行った。潜在クラス分析は、考察している確率的現象を、複数の単純な確率分布の重ね合わせとして構成される「潜在クラスモデル」によって捉える手法である。潜在クラス分析により、各仮想事例において、20個の必須アセスメント項目全体を通して現れる回答傾向を、数値的探索により推定される潜在クラスモデルの個々の構成要素である確率分布(潜在クラスと呼ばれる)として抽出することができる。本分析ではさらに、アセスメント結果を作成した職員の組織区分(児童相談所か市区町村か)および経験年数区分(5年以上か5年未満か)の二つの属性が、本分析で抽出された回答傾向とどのように関係するのかについても考察する。

仮想事例ごとに推定された潜在クラスモデルが複数の潜在クラスから構成される場合、それは回答パターンに複数の異なる傾向が検出されたことを意味する。それらの異なる回答傾向はアセスメントツールの評定者間一致性を阻害する要因として捉えられる。そのような評定者間一致性の阻害要因を検出して、組織区分と経験年数区分の観点から考察することが、本分析の目的である。

以下に、本章で示す分析結果の数値の意味を理解するのに必要かつ十分な知識として、本分析の技術的な詳細を与えておく。

アセスメント項目の個数20を M とおき、正の整数 T を潜在クラスモデルの構成要素の個数とすると、本分析で考察する潜在クラスモデルは以下の構造から成る確率モデルである：

- ・ $1 + M$ 個の確率変数： 潜在変数と呼ばれる確率変数 C 、ならびに M 個の必須アセスメント項目に対応する M 個の確率変数 X_1, X_2, \dots, X_M を考える。潜在変数 C は $1, 2, \dots, T$ のいずれかの値を取る。各 X_i は顕在変数と呼ばれ、1:該当、2:非該当、3:不明、の三つのカテゴリ値のいずれかを取る。

- ・ $T + 3TM$ 個の確率パラメータ： 各 $k = 1, 2, \dots, T$ に対し、潜在変数 C が値 k を取る確率 $p_k = \Pr(C = k)$ 、さらに各 $i = 1, 2, \dots, M$ 及び各 $a = 1, 2, 3$ に対して $C = k$ の下で X_i が値 a を取る条件付き確率 $q_k(i, a) = \Pr(X_i = a | C = k)$ をパラメータとして考える。

- ・ 確率分布： 基礎となる確率の値を定める上述のパラメータに加えて、各 $k = 1, 2, \dots, T$ に対し、 $C = k$ の下での条件付き分布において X_1, X_2, \dots, X_M は独立であるとの仮定を導入することで確率分布を定める。すなわち、 $C = k, X_1 = a_1, \dots, X_M = a_M$ となる確率を次式で定義する： $\Pr(C = k, X_1 = a_1, \dots, X_M = a_M) = p_k q_k(1, a_1) \cdots q_k(M, a_M)$ 。以上が、潜在クラスモデルの定義である。

各仮想事例における潜在クラスの推定は最尤法により行う。すなわち、サンプルが抽出された母集団の分布が上記で定義したパラメータ付きの潜在クラスモデルであると仮定した時、その母集団から本調査で得られたサンプルの集合が生じる確率(尤度) L を最大化するようなパラメータ値を推定する。

次に、サンプルの「帰属クラス」を以下のように定義する。各仮想事例に対して推定された潜在クラスモデルにおいて、個々のサンプル $a = (a_1, a_2, \dots, a_M)$ の潜在クラス k への帰属確率は次のように表される： $\Pr(C = k | X_1 = a_1, \dots, X_M = a_M)$ 。この条件付き確率を最大にする k の値が示す潜在クラスを、サンプル a の帰属クラスと呼び、またサンプル a は潜在クラス k に帰属すると言う。本分析では、各潜在クラスに対して、その潜在クラスに帰属するサンプルの有する属性（組織区分および経験年数区分）の偏りを分析することで、その潜在クラスの示す回答傾向の特徴を考察する。

最後に、クラスの個数を定めるパラメータ T の値を決定する方法について述べる。一般に、潜在クラスモデルを構成するクラスの個数 T の値を大きくするほど、より広範囲の潜在クラスモデルを探索できるようになり、より大きな尤度 L を実現できるようになる。しかし、クラスの個数 T が多くなり過ぎると、推定されるモデルは、サンプルに現れるノイズに過剰に適合することとなり、真の母集団分布からは遠ざかる。したがって、母集団分布をより適切に推定するためには、一定水準の適合の良さが認められる範囲内で、できる限り小さなクラス数 T を選択することが重要である。そのため、クラス数 T の選択基準がこれま

でに幾つか提案されている。それらの選択基準のうち、本分析では、次の式で定義されるBIC (Bayesian Information Criterion, Schwarzら, 1978) と呼ばれる値を最小化するクラス数 T を選択するという基準を採用した: $BIC = -2 \log \hat{L} + \alpha \log N$. ただし、 N はサンプルの個数、 \hat{L} は最尤法によって得られる尤度 L の最大値、 α は前述したパラメータの個数である。なお、以上で述べた潜在クラスモデルの推定は、統計解析環境Rのパッケージ `poLCA v 1.4.1` (Linzerら, 2011) を利用して行った。

6.2.3.2 傾向スコアマッチング

上述の潜在クラス分析により、本研究では経験年数区分等の区分の違いと20個の必須アセスメント項目全体の回答傾向の関係を抽出した。本研究ではさらに、区分の違いによる影響を受けやすい項目と受けにくい項目はそれぞれどのようなものかを考察することを目的とした分析を行う。具体的には、20個の必須アセスメント項目のそれぞれに対し経験年数区分の違いが回答傾向にどのような影響を与えるのかを傾向スコアマッチングを用いて分析した。

経験年数区分といったような特定の区分に着目し、割り当てられた区分の違いが結果に与える純粋な影響について調査を行う場合には、例えば組織区分の違いや管轄している地域の違いなど、その他の変数が結果に与える影響を除く必要がある。この方法の一つとして、Rosenbaum & Rubin(1983)は傾向スコア (Propensity score) 法を提案した。この傾向スコア法では複数存在するその他の変数 (共変量) について、まずそれを傾向スコアという一つの値に集約することを行う。そして区分内のそれぞれの群からその値が近いペアを作成 (マッチング) していくことを行う。これにより、区分のみが異なり、その他の変数については傾向が似ているペアのデータが作成できる。このデータを用いて、区分の違いが結果変数に与える純粋な影響の調査を本研究では行う。

以下に傾向スコア算出に関する技術的な詳細を示す。まず i 番目のデータ点における共変量ベクトルを x_i 、区分内の群1と群2のどちらの群に割り当てられているかを表す割り当て変数を z_i としたとき x_i が群1に割り当てられる確率 e_i は以下のように表される (式6.4)。この e_i が i 番目のデータ点における傾向スコアとして定義される (Rosenbaum & Rubin, 1983)。

$$e_i = \Pr(z_i = 1 | x_i) \quad (\text{式6.4})$$

この傾向スコアの上で、経験年数5年未満の群と5年以上の群それぞれから値の類似するペアを抽出 (マッチング) することで、共変量の影響を除いた、経験年数区分の違いがアセスメント項目の評価に与える純粋な影響を調査する。この際、マッチングにより2群間の共変量のバランスがどれほど調整されたのかを示す指標として、以下の式で示される各共変量の標準化平均差 (SMD) が存在する (式6.5)。

$$SMD = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{(s_1^2 + s_2^2)/2}} \quad (\text{式6.5})$$

この標準化平均差は各共変量について計算され、 \bar{X}_1 、 \bar{X}_2 はそれぞれ群1と群2におけるその共変量の平均値、 s_1^2 、 s_2^2 はそれぞれ群1と群2におけるその共変量の分散を表す。その共変

量において2群間でバランスが取れているかを確認する上では、その共変量の標準化平均差が0.1以下であることが推奨される (zhang et al., 2019)。

本研究の分析では、まず傾向スコア算出の際にはロジスティック回帰を行い、そうして得られた傾向スコアの上で、 k 最近傍法により経験年数5年未満と5年以上の群それぞれから値が近いペアを、ペア間の傾向スコアの差が小さい順に抽出していく。そしてどちらかの群の全てのデータがマッチングに用いられるか、マッチングした際のペア間の傾向スコアの差が0.2以上になるまでこの作業を繰り返す。こうして得られたデータセットに対し、各群における各共変量の標準化平均差を算出することで共変量のバランスが調整されていることを確認する。その後、20個の必須アセスメント項目に対し、区分の違いにより回答傾向が異なっているかどうかを項目別・仮想事例別に自由度2、有意水準5%でのカイ2乗検定を用いて検定する。

6.3 結果

6.3.1 回収率・欠損・除外データ

3.3.2の手順に基づいてデータを収集した結果、下表のような回収率となった(表6.3.1)。調査全体の回収率同様3割程度の回収率である。本パートに特筆すべき欠損は確認されなかったが、他パートと同様に回答不良者(同一回答)と考えられる回答組織を4件除外している。

表6.3.1 送付数・回収数・回収率

事例番号	児童相談所	送付数			回収数			回収率		
		市区町村	各グループ合計	児童相談所	市区町村	各グループ合計	児童相談所	市区町村	各グループ合計	
回答グループ1	1	36	316	352	27	80	107	0.75	0.25	0.30
回答グループ1	2	36	316	352	26	67	93	0.72	0.21	0.26
回答グループ1	3	36	316	352	24	65	89	0.67	0.21	0.25
回答グループ2	3	36	317	353	23	73	96	0.64	0.23	0.27
回答グループ2	4	36	317	353	23	62	85	0.64	0.20	0.24
回答グループ2	5	36	317	353	22	62	84	0.61	0.20	0.24
回答グループ3	5	35	316	351	10	80	90	0.29	0.25	0.26
回答グループ3	6	35	316	351	8	72	80	0.23	0.23	0.23
回答グループ3	7	35	316	351	8	71	79	0.23	0.22	0.23
回答グループ4	7	36	316	352	22	71	93	0.61	0.22	0.26
回答グループ4	8	36	316	352	20	63	83	0.56	0.20	0.24
回答グループ4	9	36	316	352	20	62	82	0.56	0.20	0.23
回答グループ5	9	36	316	352	20	84	104	0.56	0.27	0.30
回答グループ5	10	36	316	352	20	73	93	0.56	0.23	0.26
回答グループ5	11	36	316	352	19	75	94	0.53	0.24	0.27
回答グループ6	11	35	314	350	16	63	79	0.46	0.20	0.23
回答グループ6	12	35	314	350	16	56	72	0.46	0.18	0.21
回答グループ6	1	35	314	350	16	58	74	0.46	0.18	0.21

6.3.2 セーフティアセスメント項目の評定一致率

紙幅の都合により、単純集計による評定一致率をまとめた表は巻末資料とした。組織(児相・市区町村)別の単純集計による評定一致率の特徴については、以下の5パターンに分けられた(下表数値は、評定一致率=(各回答数/総回答数)×100を表す)。

(1) 児相・市区町村が一貫した回答をしているパターン

typeID		1_該当	2_非該当	3_不明・情報未取得
支援/介入の困難/資源不足/				
1	児相	75.0	12.5	12.5
1	市区町村	79.2	5.6	15.3

(2) 児相・市区町村ともに、非該当または不明・情報未取得の判断に困っている(と想定される)もしくは当該判断が一貫しない回答パターン

養育者に対する挑発/エスカレートする行為/				
11	市区町村	1.4	50.0	48.6
11	児相	0.0	62.5	37.5

(3) 児相・市区町村ともに、該当または非該当の判断に困っている(と想定される)もしくは当該判断が一貫しない回答パターン

児童および養育者の居所が不明/				
128	児相	56.2	43.8	0.0
128	市区町村	49.2	47.6	3.2

(4) 児相・市区町村ともに、どの選択肢に回答すべきか困っている(と想定される)もしくは組織間で回答が一貫しない回答パターン

養育者の家庭外ストレス/				
14	児相	31.2	37.5	31.2
14	市区町村	33.3	31.7	34.9

(5) 児相と市区町村で、判断が分かれる回答パターン

不適切な養育環境/				
19	児相	56.5	4.3	39.1
19	市区町村	38.4	5.5	56.2

そして、クリッペンドルフの α 係数を事例別、組織別、および3回提示される仮想事例について回答者の割り当てられた回答グループごとにその仮想事例の提示の順番が異なるためそのグループ(組織・事例・提示順)別に算出し、表6.3.2および表6.3.3にまとめた。全区分を通して、各項目のクリッペンドルフの α 係数が0.27から0.43程度の範囲内の値を取った。本指標がどれくらいの値をとれば、信頼性の有無について言及できるかについて具体的な基準は存在しないが、いずれの項目も決して高い信頼性を有しているとはいえないだろう。一方で、一部項目で特筆して信頼性が高い、もしくは低い項目は存在していないと考えることができるので、様々な観点から信頼性向上への糸口を見つけることが可能かもしれない。ここ

での様々な観点として考えられることは、(1)測定の問題、(2)評定者の問題、(3)ほかの要因による影響等の理由が考えられるだろう。

まず、(1)測定の問題として、本調査で用いた場面想定法もしくは、その調査内の仮想事例がこちらの意図したように作用していなかった場合(実際のセーフティアセスメント項目の評価プロセスと異なる形で回答されていた場合)、下記の信頼性が、真の信頼性よりも過小・過大に算出されている可能性がある。(2)評定者の問題として、まず児童虐待対応に関するアセスメントは性格特性を回答するような尺度とは異なり、初心者が評定を行うことが難しいことが挙げられる。また、熟練者であっても事例内容によっては判断が揺らぐ可能性がある。このことから、特定の項目が提示されたときにどのように判断することが子供の安全を守るリスク評価を行う上で妥当かについて職員間および組織間で議論や情報共有を積極的に行う必要がある。しかしながら、このような取り組みを行っている(行える)組織はとても多いとはいえない状況であるため、セミナーや研修会といったリスクを評価するための技術習得の場を積極的に設けることによって項目の信頼性を向上させることができる可能性がある。(3)ほかの要因による影響として、(2)の内容と重複するが、回答者の経験値によってセーフティアセスメント項目の評価に差が生まれ、項目の信頼性を低下させている可能性がある。もし、この要因を特定し、そこに適切な介入を行うことができれば、項目の信頼性を向上させることができるかもしれない。これについて、次項にて、回答者の経験年数という要因が、項目評価に影響を与えているのかについて詳細に検討していく。

表6.3.2 事例別に算出したクリッペンドルフの α 係数

表 事例別に算出したクリッペンドルフの α 係数					
組織	事例番号	提示順(13={1,3})	項目数	評価者数	Krippendorff's α
共通	1	13	57	181	0.346
共通	2	2	39	93	0.396
共通	3	13	58	185	0.295
共通	4	2	39	85	0.345
共通	5	13	57	174	0.368
共通	6	2	38	80	0.311
共通	7	13	56	172	0.322
共通	8	2	38	83	0.346
共通	9	13	57	186	0.328
共通	10	2	39	93	0.283
共通	11	13	57	173	0.374
共通	12	2	38	72	0.294

表6.3.3 組織および回答グループ別に算出したクリッペンドルフの α 係数

事例番号	提示順	項目数	評価者数			Krippendorff's α		
			児童相談所	市区町村	共通	児童相談所	市区町村	共通
1	1	38	27	80	107	0.384	0.362	0.368
2	2	38	26	67	93	0.386	0.398	0.396
3	3	38	24	65	89	0.289	0.306	0.300
3	1	38	23	73	96	0.315	0.280	0.289
4	2	38	23	62	85	0.356	0.344	0.345
5	3	38	22	62	84	0.390	0.329	0.342
5	1	37	10	80	90	0.417	0.390	0.394
6	2	37	8	72	80	0.318	0.308	0.311
7	3	37	8	71	79	0.281	0.294	0.297
7	1	37	22	71	93	0.383	0.323	0.336
8	2	37	20	63	83	0.389	0.335	0.346
9	3	37	20	62	82	0.390	0.326	0.337
9	1	38	20	84	104	0.346	0.305	0.311
10	2	38	20	73	93	0.290	0.289	0.283
11	3	38	19	75	94	0.427	0.376	0.385
11	1	37	16	63	79	0.399	0.354	0.363
12	2	37	16	56	72	0.295	0.293	0.294
1	3	37	16	58	74	0.310	0.312	0.313

6.3.3 回答者属性による回答傾向の比較結果（潜在クラス分析）

本項では、12仮想事例の回答データに対する潜在クラス分析の結果を報告する。6.2.3.1で説明したように、本分析では各仮想事例に対して潜在クラスモデルを推定した。また、組織区分ごと、経験年数区分ごとの回答が、各潜在クラスへ帰属する割合について分析した。

6.3.3.1 クラス数の決定

1から10までのクラス数を設定して潜在クラスモデルの推定を行ったところ、仮想事例8においてはクラス数1の時に、仮想事例2、3、4、5、6、7、10、12においてはクラス数2の時に、仮想事例1、9、11においてはクラス数3の時に、BICが最小の値を取った。結果を図6.3.3.1(1)-(2)に示す。各仮想事例に対して推定した潜在クラスモデルのクラス数は、以上の結果に基づき決定した。

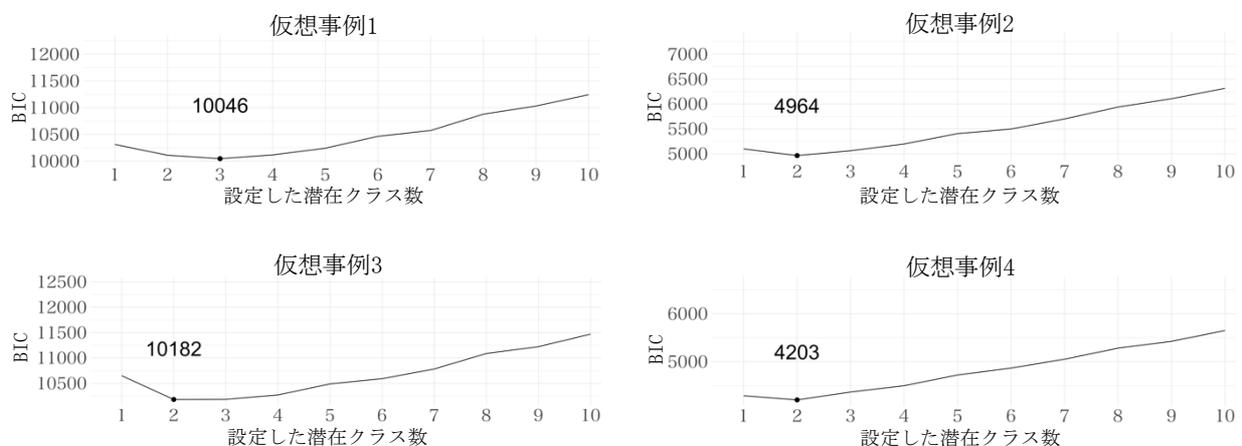


図6.3.3.1 (1) 各仮想事例に対して推定された潜在クラスモデルのクラス数別のBIC

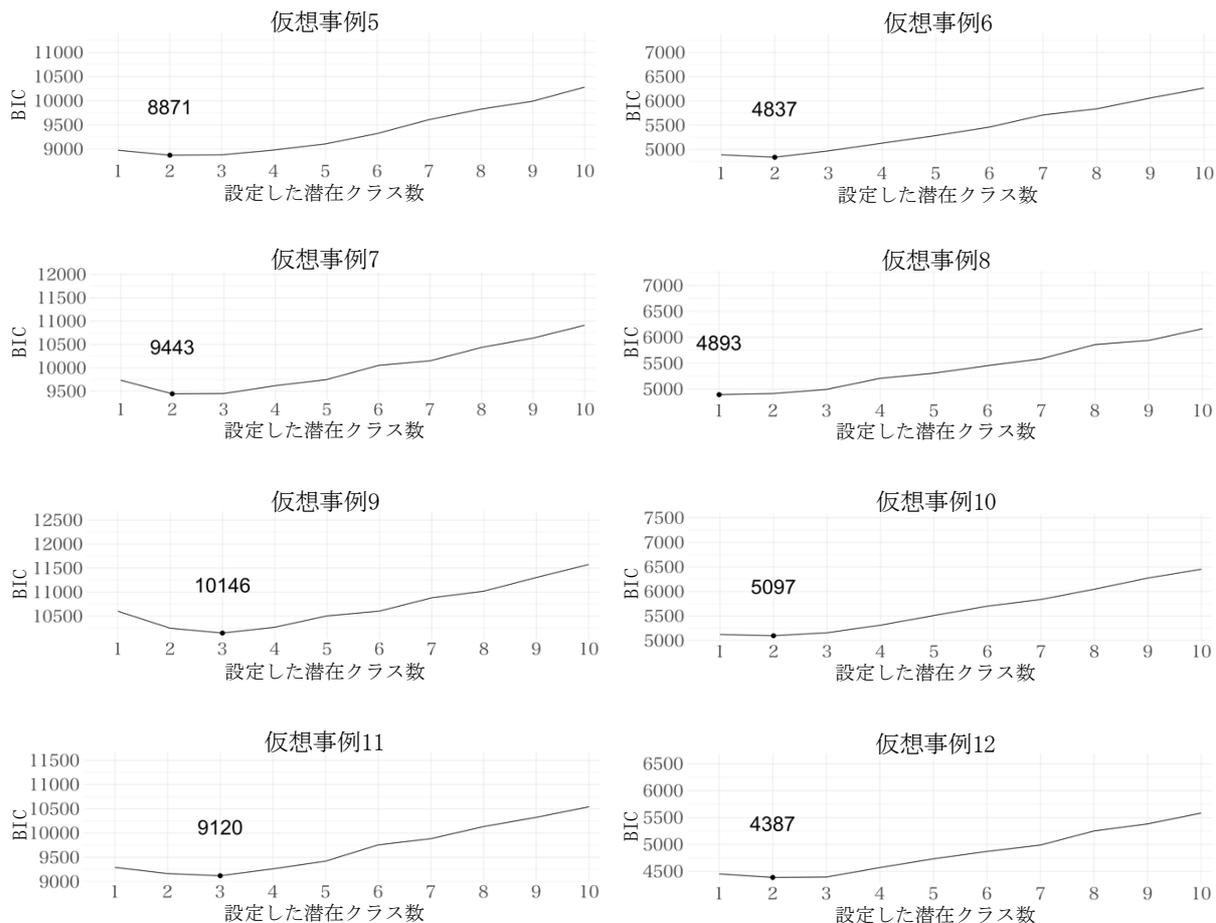


図6.3.3.1 (2) 各仮想事例に対して推定された潜在クラスモデルのクラス数別のBIC

6.3.3.2 推定された潜在クラスモデル

各仮想事例に対して潜在クラスモデルの推定を行った。その結果として得られた各パラメータの推定値を表6.3.3.1に示す。

クラス数が2となった潜在クラスモデルは、そのすべてが、20個の項目全体にわたって「不明」と回答しやすい傾向を持つクラスと、それ以外のクラスに分かれた。表6.3.3.1に示す通り、クラス数が2である仮想事例2、3、4、5、6、7、10、12においては、クラス2がクラス1と比べて、20個のアセスメント項目のほぼ全てにおいて「不明」と回答する確率が高くなっている。本項では以下、この「不明」と回答する確率が高いほうのクラスを「不明指向クラス」と言う。

クラス数が3となった仮想事例1、9、11においても、クラス数が2の場合と同様に、その潜在クラスモデルのすべてが、ほぼ全てのアセスメント項目に関して他のクラスよりも「不明」と回答する確率が高いクラスを有していた。さらに、同様の意味で、「該当」と回答する確率が高いクラス、「非該当」と回答する確率が高いクラスを有している。本項では以下、これらのクラスを順に、「不明指向クラス」「該当指向クラス」「非該当指向クラス」と言う。表6.3.3.1の中では、仮想事例1、9、11のクラス1、クラス2、クラス3が、それぞれ該当指向クラス、非該当指向クラス、不明クラスである。

表6.3.3.1 推定された潜在クラスモデル

	仮想事例1									仮想事例2					
	クラス1($p_1 = 0.25$)			クラス2($p_2 = 0.30$)			クラス3($p_3 = 0.45$)			クラス1($p_1 = 0.33$)			クラス2($p_2 = 0.67$)		
	該当	非該当	不明	該当	非該当	不明	該当	非該当	不明	該当	非該当	不明	該当	非該当	不明
項目1	0.78	0.16	0.07	0.55	0.32	0.13	0.62	0.07	0.30	0.85	0.07	0.05	0.51	0.14	0.32
項目2	1.00	0.00	0.00	0.51	0.13	0.35	0.41	0.08	0.50	0.87	0.00	0.13	0.90	0.00	0.10
項目3	0.69	0.11	0.20	0.29	0.36	0.33	0.36	0.07	0.56	0.74	0.07	0.16	0.93	0.00	0.07
項目4	0.91	0.05	0.04	0.86	0.07	0.07	0.88	0.01	0.11	0.23	0.47	0.27	0.34	0.11	0.56
項目5	0.72	0.10	0.17	0.38	0.17	0.45	0.31	0.09	0.58	0.90	0.03	0.07	0.86	0.02	0.13
項目6	0.53	0.06	0.40	0.22	0.35	0.41	0.17	0.08	0.75	0.60	0.30	0.10	0.65	0.10	0.25
項目7	0.69	0.00	0.31	0.63	0.13	0.24	0.39	0.01	0.60	0.71	0.03	0.26	0.66	0.02	0.32
項目8	0.83	0.04	0.13	0.71	0.18	0.09	0.75	0.10	0.16	0.03	0.87	0.07	0.00	0.65	0.33
項目9	1.00	0.00	0.00	0.82	0.09	0.09	0.87	0.01	0.12	0.77	0.10	0.13	0.87	0.00	0.13
項目10	0.98	0.00	0.02	0.98	0.00	0.02	0.93	0.00	0.07	0.10	0.43	0.45	0.13	0.08	0.79
項目11	0.79	0.02	0.18	0.36	0.56	0.06	0.43	0.17	0.41	0.13	0.64	0.16	0.03	0.25	0.72
項目12	0.89	0.07	0.03	0.34	0.57	0.08	0.34	0.14	0.52	0.90	0.03	0.07	0.84	0.00	0.16
項目13	0.69	0.00	0.31	0.42	0.28	0.28	0.28	0.10	0.62	0.07	0.43	0.47	0.06	0.02	0.92
項目14	1.00	0.00	0.00	0.68	0.17	0.15	0.64	0.04	0.31	0.16	0.64	0.20	0.02	0.30	0.69
項目15	0.80	0.06	0.14	0.53	0.21	0.22	0.31	0.03	0.65	0.70	0.23	0.03	0.71	0.03	0.25
項目16	1.00	0.00	0.00	0.90	0.04	0.07	0.84	0.01	0.15	0.60	0.23	0.17	0.54	0.24	0.21
項目17	0.00	0.83	0.16	0.00	0.88	0.12	0.00	0.70	0.28	0.00	0.93	0.03	0.00	0.54	0.46
項目18	0.00	0.30	0.68	0.00	0.73	0.24	0.00	0.02	0.97	0.00	0.79	0.18	0.00	0.16	0.84
項目19	0.23	0.05	0.72	0.02	0.37	0.59	0.08	0.01	0.89	0.00	0.70	0.26	0.10	0.09	0.81
項目20	0.91	0.00	0.09	0.56	0.20	0.24	0.51	0.00	0.47	0.03	0.56	0.37	0.00	0.11	0.89

	仮想事例3						仮想事例4					
	クラス1($p_1 = 0.41$)			クラス2($p_2 = 0.59$)			クラス1($p_1 = 0.56$)			クラス2($p_2 = 0.44$)		
	該当	非該当	不明									
項目1	0.25	0.44	0.30	0.32	0.23	0.45	0.79	0.15	0.06	0.57	0.08	0.35
項目2	0.14	0.43	0.41	0.16	0.09	0.75	0.31	0.21	0.48	0.41	0.00	0.57
項目3	0.47	0.21	0.31	0.42	0.02	0.56	0.90	0.02	0.08	0.92	0.00	0.08
項目4	0.53	0.30	0.16	0.46	0.06	0.47	0.92	0.04	0.04	0.76	0.00	0.24
項目5	0.39	0.07	0.53	0.12	0.01	0.87	0.71	0.02	0.25	0.54	0.00	0.46
項目6	0.83	0.10	0.07	0.77	0.00	0.21	1.00	0.00	0.00	0.95	0.00	0.05
項目7	0.24	0.21	0.54	0.29	0.01	0.69	0.15	0.15	0.71	0.16	0.03	0.81
項目8	0.01	0.74	0.23	0.02	0.43	0.55	0.10	0.56	0.33	0.05	0.14	0.81
項目9	0.13	0.22	0.63	0.05	0.03	0.92	0.29	0.15	0.56	0.22	0.03	0.76
項目10	0.07	0.25	0.66	0.03	0.03	0.93	0.02	0.27	0.71	0.03	0.00	0.97
項目11	0.00	0.84	0.13	0.00	0.35	0.65	0.00	1.00	0.00	0.00	0.00	0.97
項目12	0.12	0.39	0.46	0.12	0.07	0.81	0.69	0.06	0.25	0.76	0.03	0.22
項目13	0.32	0.30	0.37	0.17	0.03	0.80	0.92	0.02	0.06	0.92	0.00	0.08
項目14	0.00	0.64	0.34	0.02	0.19	0.79	0.08	0.29	0.62	0.05	0.03	0.92
項目15	0.07	0.84	0.08	0.05	0.62	0.34	0.60	0.06	0.33	0.59	0.00	0.38
項目16	0.93	0.01	0.05	0.93	0.02	0.05	0.92	0.02	0.06	0.89	0.00	0.11
項目17	0.00	0.83	0.15	0.03	0.42	0.55	0.67	0.27	0.04	0.70	0.05	0.24
項目18	0.00	0.68	0.29	0.01	0.12	0.85	0.33	0.38	0.29	0.08	0.16	0.76
項目19	0.44	0.16	0.38	0.41	0.02	0.57	0.96	0.02	0.02	1.00	0.00	0.00
項目20	0.00	0.55	0.43	0.01	0.09	0.90	0.31	0.21	0.48	0.30	0.00	0.70

	仮想事例5						仮想事例6					
	クラス1($p_1 = 0.54$)			クラス2($p_2 = 0.46$)			クラス1($p_1 = 0.47$)			クラス2($p_2 = 0.53$)		
	該当	非該当	不明	該当	非該当	不明	該当	非該当	不明	該当	非該当	不明
項目1	0.90	0.06	0.04	0.56	0.10	0.35	0.81	0.11	0.08	0.77	0.02	0.21
項目2	0.69	0.15	0.17	0.66	0.07	0.26	0.13	0.52	0.35	0.28	0.30	0.42
項目3	0.85	0.08	0.07	0.67	0.00	0.32	0.35	0.34	0.31	0.47	0.03	0.48
項目4	0.08	0.12	0.80	0.06	0.00	0.92	0.00	0.38	0.62	0.05	0.14	0.81
項目5	0.96	0.00	0.04	0.84	0.01	0.12	0.78	0.05	0.16	0.61	0.05	0.35
項目6	1.00	0.00	0.00	0.94	0.00	0.05	0.95	0.00	0.05	0.93	0.00	0.07
項目7	0.73	0.05	0.22	0.62	0.01	0.36	0.49	0.32	0.19	0.62	0.05	0.33
項目8	0.02	0.48	0.48	0.00	0.12	0.86	0.11	0.54	0.35	0.14	0.14	0.72
項目9	0.85	0.03	0.11	0.56	0.01	0.42	0.24	0.30	0.44	0.26	0.07	0.67
項目10	0.88	0.01	0.09	0.78	0.01	0.20	0.24	0.38	0.38	0.09	0.04	0.87
項目11	0.04	0.60	0.34	0.04	0.10	0.85	0.00	0.73	0.27	0.02	0.32	0.66
項目12	0.96	0.03	0.00	0.75	0.02	0.21	0.78	0.13	0.08	0.84	0.05	0.09
項目13	0.55	0.08	0.36	0.26	0.05	0.66	0.89	0.00	0.11	0.95	0.00	0.05
項目14	0.06	0.26	0.67	0.03	0.04	0.91	0.16	0.32	0.52	0.16	0.00	0.84

項目15	0.56	0.11	0.32	0.36	0.02	0.61	0.11	0.51	0.38	0.12	0.19	0.70
項目16	0.52	0.15	0.32	0.31	0.07	0.60	0.05	0.61	0.33	0.19	0.05	0.76
項目17	0.48	0.32	0.19	0.22	0.32	0.43	0.46	0.52	0.03	0.65	0.21	0.14
項目18	0.73	0.12	0.15	0.51	0.09	0.39	0.05	0.62	0.33	0.00	0.00	1.00
項目19	0.99	0.00	0.00	0.93	0.00	0.06	0.57	0.19	0.24	0.60	0.00	0.40
項目20	0.47	0.27	0.26	0.20	0.30	0.47	0.54	0.22	0.24	0.42	0.09	0.49

		仮想事例7						仮想事例8		
		クラス1($p_1 = 0.61$)			クラス2($p_2 = 0.39$)			クラス1($p_1 = 1.00$)		
		該当	非該当	不明	該当	非該当	不明	該当	非該当	不明
項目1		0.86	0.08	0.06	0.86	0.00	0.14	0.60	0.27	0.13
項目2		0.52	0.26	0.22	0.27	0.13	0.60	0.86	0.05	0.10
項目3		0.98	0.02	0.00	0.99	0.00	0.01	0.43	0.14	0.41
項目4		0.10	0.38	0.52	0.02	0.11	0.87	0.98	0.00	0.01
項目5		0.75	0.01	0.24	0.34	0.00	0.66	0.48	0.18	0.34
項目6		0.86	0.04	0.11	0.78	0.00	0.22	0.42	0.17	0.41
項目7		0.37	0.16	0.46	0.16	0.01	0.83	0.98	0.00	0.01
項目8		0.24	0.61	0.15	0.21	0.25	0.54	0.43	0.39	0.18
項目9		0.29	0.16	0.54	0.06	0.03	0.92	0.67	0.01	0.31
項目10		0.16	0.41	0.41	0.04	0.08	0.86	0.64	0.35	0.00
項目11		0.08	0.70	0.22	0.01	0.37	0.62	0.70	0.05	0.25
項目12		0.52	0.22	0.26	0.01	0.04	0.94	0.57	0.25	0.18
項目13		0.69	0.11	0.20	0.51	0.00	0.49	0.51	0.11	0.39
項目14		0.23	0.38	0.40	0.13	0.13	0.74	0.02	0.54	0.43
項目15		0.98	0.02	0.00	0.99	0.00	0.01	0.61	0.14	0.24
項目16		0.59	0.13	0.27	0.32	0.03	0.66	0.11	0.39	0.51
項目17		0.05	0.66	0.28	0.03	0.30	0.67	0.02	0.89	0.05
項目18		0.86	0.08	0.05	0.84	0.02	0.12	0.00	0.53	0.46
項目19		0.65	0.09	0.26	0.59	0.00	0.41	0.19	0.30	0.51
項目20		0.63	0.07	0.30	0.51	0.00	0.49	0.49	0.18	0.31

		仮想事例9						仮想事例10								
		クラス1($p_1 = 0.36$)			クラス2($p_2 = 0.26$)			クラス3($p_3 = 0.38$)			クラス1($p_1 = 0.42$)			クラス2($p_2 = 0.58$)		
		該当	非該当	不明	該当	非該当	不明	該当	非該当	不明	該当	非該当	不明	該当	非該当	不明
項目1		0.77	0.11	0.12	0.74	0.15	0.08	0.70	0.04	0.24	0.76	0.20	0.05	0.70	0.02	0.28
項目2		0.38	0.11	0.51	0.34	0.45	0.20	0.18	0.04	0.76	0.90	0.08	0.00	0.70	0.00	0.30
項目3		0.96	0.00	0.04	0.68	0.17	0.15	0.57	0.03	0.39	0.92	0.03	0.05	0.93	0.00	0.07
項目4		0.90	0.01	0.08	0.44	0.36	0.17	0.60	0.01	0.37	0.64	0.13	0.23	0.47	0.15	0.38
項目5		0.89	0.00	0.11	0.74	0.06	0.17	0.44	0.01	0.54	0.89	0.00	0.08	0.71	0.00	0.29
項目6		0.98	0.00	0.02	0.52	0.21	0.23	0.53	0.03	0.43	1.00	0.00	0.00	0.94	0.02	0.04
項目7		0.64	0.01	0.35	0.62	0.12	0.26	0.34	0.00	0.66	0.77	0.03	0.20	0.83	0.02	0.15
項目8		0.10	0.79	0.11	0.00	0.98	0.02	0.03	0.61	0.36	0.21	0.55	0.24	0.23	0.34	0.43
項目9		0.39	0.10	0.51	0.19	0.23	0.58	0.06	0.00	0.91	0.90	0.05	0.05	0.85	0.02	0.13
項目10		0.48	0.06	0.46	0.20	0.39	0.38	0.20	0.04	0.75	0.64	0.23	0.13	0.62	0.11	0.26
項目11		0.52	0.19	0.29	0.22	0.38	0.38	0.25	0.06	0.68	0.13	0.77	0.10	0.11	0.61	0.28
項目12		0.99	0.00	0.00	0.85	0.08	0.06	0.68	0.06	0.27	0.81	0.03	0.16	0.64	0.04	0.32
項目13		1.00	0.00	0.00	0.96	0.00	0.04	0.93	0.00	0.07	0.41	0.18	0.41	0.16	0.07	0.77
項目14		0.10	0.03	0.87	0.02	0.59	0.39	0.04	0.04	0.92	0.10	0.22	0.65	0.02	0.02	0.96
項目15		0.56	0.05	0.39	0.32	0.34	0.34	0.21	0.07	0.71	0.65	0.20	0.15	0.41	0.04	0.55
項目16		0.33	0.03	0.64	0.41	0.13	0.44	0.10	0.00	0.90	0.18	0.18	0.64	0.09	0.00	0.91
項目17		0.01	0.87	0.10	0.00	0.92	0.04	0.01	0.75	0.21	0.13	0.74	0.13	0.07	0.12	0.81
項目18		0.03	0.43	0.54	0.00	0.98	0.02	0.00	0.17	0.83	0.15	0.63	0.22	0.13	0.04	0.83
項目19		0.87	0.08	0.05	0.61	0.20	0.19	0.45	0.03	0.51	0.51	0.15	0.34	0.48	0.00	0.52
項目20		0.79	0.03	0.18	0.42	0.42	0.16	0.43	0.07	0.49	0.12	0.25	0.62	0.02	0.05	0.93

		仮想事例11						仮想事例12								
		クラス1($p_1 = 0.37$)			クラス2($p_2 = 0.42$)			クラス3($p_3 = 0.21$)			クラス1($p_1 = 0.53$)			クラス2($p_2 = 0.47$)		
		該当	非該当	不明	該当	非該当	不明	該当	非該当	不明	該当	非該当	不明	該当	非該当	不明
項目1		0.99	0.00	0.01	0.86	0.07	0.07	0.71	0.07	0.22	0.80	0.08	0.13	0.63	0.03	0.34
項目2		0.63	0.09	0.27	0.26	0.37	0.37	0.39	0.15	0.46	0.87	0.00	0.13	0.85	0.06	0.09
項目3		0.75	0.00	0.24	0.57	0.28	0.14	0.44	0.00	0.56	0.90	0.05	0.05	0.66	0.03	0.32
項目4		0.08	0.03	0.89	0.03	0.31	0.67	0.00	0.00	1.00	0.20	0.59	0.21	0.11	0.09	0.80
項目5		0.71	0.02	0.27	0.45	0.04	0.49	0.25	0.00	0.75	0.80	0.05	0.15	0.80	0.00	0.20
項目6		1.00	0.00	0.00	0.92	0.01	0.07	0.81	0.00	0.19	0.39	0.43	0.18	0.17	0.03	0.80
項目7		0.51	0.02	0.47	0.26	0.26	0.48	0.08	0.08	0.83	0.90	0.00	0.10	0.71	0.00	0.29
項目8		0.02	0.59	0.39	0.01	0.94	0.04	0.00	0.59	0.41	0.39	0.26	0.36	0.31	0.20	0.49
項目9		0.98	0.00	0.02	0.81	0.08	0.11	0.30	0.00	0.70	0.97	0.00	0.03	0.91	0.00	0.06

項目10	0.70	0.00	0.30	0.60	0.16	0.24	0.44	0.03	0.53	0.15	0.45	0.39	0.06	0.18	0.76
項目11	0.03	0.39	0.57	0.01	0.94	0.04	0.03	0.49	0.48	0.20	0.77	0.03	0.03	0.51	0.46
項目12	0.98	0.02	0.00	0.90	0.01	0.09	0.50	0.08	0.42	0.76	0.10	0.13	0.55	0.03	0.42
項目13	0.95	0.00	0.05	0.99	0.01	0.00	0.78	0.00	0.22	0.49	0.28	0.23	0.34	0.00	0.66
項目14	0.35	0.08	0.57	0.32	0.39	0.29	0.12	0.31	0.57	0.13	0.36	0.52	0.00	0.00	1.00
項目15	1.00	0.00	0.00	0.95	0.05	0.00	0.78	0.05	0.16	0.58	0.15	0.24	0.38	0.09	0.54
項目16	0.54	0.04	0.42	0.26	0.28	0.45	0.19	0.13	0.67	0.41	0.36	0.23	0.23	0.17	0.60
項目17	0.02	0.86	0.11	0.00	0.92	0.05	0.00	0.84	0.11	0.00	0.80	0.20	0.00	0.77	0.23
項目18	0.27	0.02	0.71	0.12	0.44	0.43	0.05	0.17	0.78	0.00	0.67	0.33	0.00	0.23	0.77
項目19	0.49	0.02	0.50	0.42	0.07	0.52	0.33	0.00	0.67	0.41	0.39	0.20	0.11	0.20	0.69
項目20	0.05	0.13	0.82	0.08	0.65	0.28	0.03	0.25	0.73	0.38	0.10	0.52	0.20	0.00	0.80

6.3.3.3 組織区分ごとの潜在クラス帰属割合

組織区分の違いが潜在クラスへの帰属割合にどのように関係するかを分析した。各仮想事例において、組織区分ごとの回答が各潜在クラスへ帰属する割合を表6.3.3.2に示す。

表6.3.3.2 組織区分ごとの潜在クラス帰属数と帰属割合

	仮想事例1			仮想事例2		仮想事例3	
	1該当指向	2非該当指向	3不明指向	クラス1	2不明指向	クラス1	2不明指向
市区町村	37件(26.4%)	38件(27.1%)	65件(46.4%)	22件(32.8%)	45件(67.2%)	54件(39.1%)	84件(60.9%)
児童相談所	8件(18.6%)	16件(37.2%)	19件(44.2%)	9件(34.6%)	17件(65.4%)	23件(48.9%)	24件(51.1%)
合計	45件(25.0%)	54件(30.0%)	84件(46.0%)	31件(33.3%)	62件(66.7%)	77件(42.0%)	108件(58.0%)

	仮想事例4		仮想事例5		仮想事例6	
	クラス1	2不明指向	クラス1	2不明指向	クラス1	2不明指向
市区町村	34件(54.8%)	28件(45.2%)	76件(53.1%)	67件(46.9%)	34件(47.2%)	38件(52.8%)
児童相談所	14件(60.9%)	9件(39.1%)	19件(59.4%)	13件(40.6%)	3件(37.5%)	5件(62.5%)
合計	48件(56.5%)	37件(43.5%)	95件(54.3%)	80件(45.7%)	37件(46.0%)	43件(54.0%)

	仮想事例7		仮想事例8	仮想事例9		
	クラス1	2不明指向	クラス1	1該当指向	2非該当指向	3不明指向
市区町村	87件(60.8%)	56件(39.2%)	63件(100%)	52件(35.4%)	41件(27.9%)	54件(36.7%)
児童相談所	19件(63.3%)	11件(36.7%)	20件(100%)	15件(37.5%)	7件(17.5%)	18件(45.0%)
合計	106件(61.3%)	67件(38.7%)	83件(100%)	67件(36.0%)	48件(26.0%)	72件(39.0%)

	仮想事例10		仮想事例11			仮想事例12	
	クラス1	2不明指向	1該当指向	2非該当指向	3不明指向	クラス1	2不明指向
市区町村	30件(40.5%)	44件(59.5%)	48件(34.0%)	60件(42.6%)	33件(23.4%)	29件(50.0%)	29件(50.0%)
児童相談所	9件(45.0%)	11件(55.0%)	18件(51.4%)	14件(40.0%)	3件(8.6%)	10件(62.5%)	6件(37.5%)
合計	39件(41.5%)	55件(58.5%)	66件(38.0%)	74件(42.0%)	36件(20.0%)	39件(52.7%)	35件(47.3%)

クラス数が2となった仮想事例において、組織区分によって不明指向クラスへの帰属割合に差が生じるケースとしては、市区町村のほうが児童相談所よりも帰属割合が大きい場合と、逆に児童相談所のほうが市区町村よりも帰属割合が大きい場合の両方が見られた。市区町村のほうがより大きい帰属割合を持つケースとしては、仮想事例3、4、5、10、12があった。一方、児童相談所の方がより大きい帰属割合を持つケースとしては、仮想事例6があった。

クラス数が3となった仮想事例において、組織区分によって不明指向クラスへの帰属割合に差が生じるケースとしては、仮想事例9、11があった。仮想事例9では児童相談所のほうが市区町村よりも帰属割合が大きかったが、仮想事例11では逆に市区町村のほうが児童相談所よりも帰属割合が大きかった。また、児童相談所と市区町村のどちらが該当指向クラスへの帰属割合をより大きく持つかに関しても、仮想事例1と11では異なっていた。非該当指向クラスに関しても、仮想事例1と9で同様の違いが見られた。

6.3.3.4 組織区分・経験年数区分ごとの潜在クラス帰属割合

組織区分の違いに加えて、さらに経験年数区分の違いが潜在クラスへの帰属割合にどのように関係するかを分析した。各仮想事例における組織区分・経験年数区分ごとの回答が各潜在クラスへ帰属する割合を表6.3.3.3に示す。

表6.3.3.3 組織区分・経験年数区分ごとの潜在クラス帰属数と帰属割合

		仮想事例1			仮想事例2		仮想事例3	
		1該当指向	2非該当指向	3不明指向	クラス1	2不明指向	クラス1	2不明指向
市区町村	5年以上	5件(20.8%)	9件(37.5%)	10件(41.7%)	5件(33.3%)	10件(66.7%)	11件(31.4%)	24件(68.6%)
	5年未満	26件(30.6%)	26件(30.6%)	33件(38.8%)	16件(37.2%)	27件(62.8%)	33件(41.3%)	47件(58.8%)
	N/A	6件(19.4%)	3件(9.7%)	22件(71.0%)	1件(11.1%)	8件(88.9%)	10件(43.5%)	13件(56.5%)
児童相談所	5年以上	6件(23.1%)	9件(34.6%)	11件(42.3%)	5件(35.7%)	9件(64.3%)	12件(50.0%)	12件(50.0%)
	5年未満	2件(15.4%)	6件(46.2%)	5件(38.5%)	2件(22.2%)	7件(77.8%)	10件(55.6%)	8件(44.4%)
	N/A	0件(00.0%)	1件(25.0%)	3件(75.0%)	1件(33.3%)	2件(66.7%)	1件(20.0%)	4件(80.0%)

		仮想事例4		仮想事例5		仮想事例6	
		クラス1	2不明指向	クラス1	2不明指向	クラス1	2不明指向
市区町村	5年以上	8件(57.1%)	6件(42.9%)	20件(55.6%)	16件(44.4%)	11件(50.0%)	11件(50.0%)
	5年未満	16件(48.5%)	17件(51.5%)	37件(50.7%)	36件(49.3%)	19件(48.7%)	20件(51.3%)
	N/A	10件(66.7%)	5件(33.3%)	19件(55.9%)	15件(44.1%)	4件(36.4%)	7件(63.6%)
児童相談所	5年以上	9件(69.2%)	4件(30.8%)	13件(65.0%)	7件(35.0%)	1件(33.3%)	2件(66.7%)
	5年未満	3件(42.9%)	4件(57.1%)	5件(50.0%)	5件(50.0%)	2件(40.0%)	3件(60.0%)
	N/A	2件(66.7%)	1件(33.3%)	1件(50.0%)	1件(50.0%)	0件(00.0%)	0件(00.0%)

		仮想事例7		仮想事例8	仮想事例9		
		クラス1	2不明指向	クラス1	1該当指向	2非該当指向	3不明指向
市区町村	5年以上	21件(55.3%)	17件(44.7%)	13件(100%)	10件(37.0%)	6件(22.2%)	11件(40.7%)
	5年未満	45件(60.0%)	30件(40.0%)	37件(100%)	34件(37.4%)	25件(27.5%)	32件(35.2%)
	N/A	21件(70.0%)	9件(30.0%)	13件(100%)	8件(27.6%)	10件(34.5%)	11件(37.9%)
児童相談所	5年以上	11件(64.7%)	6件(35.3%)	13件(100%)	9件(36.0%)	4件(16.0%)	12件(48.0%)
	5年未満	8件(66.7%)	4件(33.3%)	5件(100%)	5件(41.7%)	3件(25.0%)	4件(33.3%)
	N/A	1件(50.0%)	1件(50.0%)	2件(100%)	1件(33.3%)	0件(00.0%)	2件(66.7%)

		仮想事例10		仮想事例11			仮想事例12	
		クラス1	2不明指向	1該当指向	2非該当指向	3不明指向	クラス1	2不明指向
市区町村	5年以上	3件(21.4%)	11件(78.6%)	7件(26.9%)	11件(42.3%)	8件(30.8%)	7件(53.8%)	6件(46.2%)
	5年未満	25件(54.3%)	21件(45.7%)	26件(33.8%)	33件(42.9%)	18件(23.4%)	18件(58.1%)	13件(41.9%)
	N/A	2件(14.3%)	12件(85.7%)	15件(39.5%)	16件(42.1%)	7件(18.4%)	4件(28.6%)	10件(71.4%)
児童相談所	5年以上	6件(42.9%)	8件(57.1%)	14件(56.0%)	9件(36.0%)	2件(8.0%)	6件(66.7%)	3件(33.3%)
	5年未満	1件(25.0%)	3件(75.0%)	2件(28.6%)	4件(57.1%)	1件(14.3%)	3件(60.0%)	2件(40.0%)
	N/A	2件(50.0%)	2件(50.0%)	2件(66.7%)	1件(33.3%)	0件(00.0%)	1件(50.0%)	1件(50.0%)

クラス数が2となった仮想事例において、経験年数区分の違いによって不明指向クラスへの帰属割合に大きな差が生じたのは、仮想事例3、4、5、10であった。仮想事例3では、児童相談所と市区町村の両方で、経験年数5年以上のほうが5年未満よりも不明指向クラスへの帰属割合が大きかった。一方、仮想事例4、5では、児童相談所と市区町村の両方で、経験年数5年未満のほうが5年以上よりも不明指向クラスへの帰属割合が大きかった。仮想事例10では、市区町村において、経験年数5年以上のほうが5年未満よりも不明指向クラスへの帰属割合がかなり大きかった。一方、児童相談所においては、逆に経験年数5年未満のほうが5年以上よりも不明指向クラスへの帰属割合が大きかった。

クラス数が3となった仮想事例において、経験年数区分の違いによって不明指向クラスへの帰属割合に大きな差が生じたのは、仮想事例9、11であった。仮想事例9では、児童相談所と市区町村の両方で、経験年数5年以上のほうが5年未満よりも不明指向クラスへの帰属割合が大きかった。仮想事例11においても、市区町村に関しては同様の結果であった。また、差がわずかであることを除けば、仮想事例11においても同様の結果が見られた。

クラス数が3となった仮想事例のいずれにおいても、経験年数区分の違いによって該当指向クラスまたは非該当指向クラスへの帰属割合に差が生じていた。仮想事例1、11では、市区町村に関しては、経験年数5年未満の方が5年以上よりも、該当指向クラスへの帰属割合が大きかった。一方、児童相談所に関しては、逆に経験年数5年以上のほうが5年未満よりも該当指向クラスへの帰属割合が大きかった。

以上に述べた通り、経験年数区分の違いによって各潜在クラスへの帰属割合に差が生じる仮想事例には様々な類型があった。経験年数が増えると不明指向クラスへの帰属割合が大きくなることもあれば小さくなることもあり、またその大小の向きが児童相談所と市区町村の間で相反する場合もあった。

6.4 考察（潜在クラス分析）

6.3.3.2で見たように、各仮想事例に対して推定された潜在クラスモデルには、クラス数が1となる場合、2となる場合、3となる場合があった。クラス数が2となった潜在クラスモデルは、不明指向クラスとそれ以外のクラスの二つから構成されていた。また、クラス数が3となった潜在クラスモデルは、不明指向クラス、該当指向クラス、非該当指向クラスの三つから構成されていた。これらの潜在クラスモデルは、6.2.3.1で述べた通り、20個の必須アセスメント項目全体にわたって現れる回答傾向を示すものとして捉えられる。すなわち、相対的な意味で、アセスメント項目の全体にわたって不明と回答しやすい傾向、該当と回答しやすい傾向、非該当と回答しやすい傾向が、事例によっては生じ得るということを本分析の結果は示唆している。

また、本分析では、組織区分と経験年数区分の違いが各潜在クラスへの帰属割合にどのように関係するかについても考察した。そこで得られた知見は、不明指向、該当指向、非該当指向といった回答傾向の生じやすさは、組織区分や経験年数区分の違い、また事例の内容の違いによって変わり得るということであった。その結果を踏まえると、本分析で見られた回答傾向は、職員が事例の内容に触れた際に生じるある種のバイアスとして捉えることもできるのかもしれない。本研究で構成するセーフティアセスメントツールの評定者間一致性の向上を目的として、利用手引きの設計や研修の実施などを行う際には、組織区分や経験年数の違いから生じるバイアス等の回答傾向の差も考慮して、その差を埋めるための工夫を盛り込むことが重要である。

6.5 結果（傾向スコアマッチング）

6.3節および6.4節の潜在クラス分析から、経験年数区分の違いにより同じ仮想事例でもアセスメント項目の回答結果に差が出る可能性が示唆された。一方で、潜在クラス分析では区分の違いが項目全体の回答傾向に与える影響を抽出したものであって、項目別の回答傾向については説明を与えていない。しかし実際には、経験年数区分の違いによって回答傾向に影響を受けやすい項目と受けにくい項目が存在する可能性がある。そのため、本節では項目別の評定者間一致性、つまり、経験年数区分の違いが個々の項目の回答傾向に与える影響につ

いて調査する。具体的には、傾向スコアマッチングを用いて共変量による影響を調整した上で、経験年数5年未満と5年以上による回答傾向の違いを仮想事例別・項目別に比較する。

6.5.1 使用データ

まず、Web調査表のパート1およびパート2より得たデータにおいて、本節の分析で用いる共変量を表6.5.1の通りに決定した。本節では、これらの共変量およびパート2の仮想事例に対するアセスメント項目の評価に欠損のないデータを抽出し、分析に用いる。この抽出の結果、経験年数5年未満のデータが949件、経験年数5年以上のデータが430件抽出された。このデータより、表6.5.1に示した共変量を用いて傾向スコアを算出した後、サンプルを経験年数5年未満と5年以上の群に分け、傾向スコアマッチングを実施する。その後、経験年数区分の違いにより20個の必須アセスメント項目の回答傾向が異なっていたかどうかをカイ2乗検定により検定する。

表6.5.1 本節にて用いる共変量の一覧

変数	値	定義
管轄人口	数値	2020年4月時点での管轄人口
現在対応中の対応件数	数値	現在対応中の児童虐待相談対応件数
18歳未満在住比率	数値	管轄地域の一般世帯数に対する18歳未満の在住している一般世帯数の比率
核世帯家族比率	数値	管轄地域の一般世帯数に対する1核世帯家族の在住している一般世帯数の比率
20歳未満人口割合	数値	管轄地域における20歳未満の人口割合
65歳以上人口割合	数値	管轄地域における65歳以上の人口割合
組織ダミー	0:市区町村, 1:児童相談所	回答者の所属組織
シート利用ダミー	0:リスクアセスメントシートを使用していない, 1:リスクアセスメントシートを使用している	リスクアセスメントシートの有無

6.5.2 使用データの基礎集計

まず、用いる変数の基本統計量を示す。表6.5.2は使用する変数のうち、数値データに関する基本統計量である。表6.5.3はカテゴリカルデータのデータ数であり、組織ダミー、シート利用ダミーそれぞれについて0と回答したデータ数と1と回答したデータ数を記載している。なお、どちらもデータ数は1379件である。

表6.5.2 基本統計量（数値データ）

	管轄人口	現在対応中の 対応件数	経験年数	一般世帯数 18歳未満在住_比率	一般世帯数 核世帯家族_比率	人口20歳未満 割合	人口65歳以上 割合
平均値	229391.5	212.68	4.08	0.22	0.57	0.17	0.29
標準偏差	908453.32	347.02	4.25	0.04	0.06	0.02	0.06
最小値	321	0	0	0.08	0.33	0.08	0.15
中央値	61457	77	3	0.23	0.57	0.17	0.28
最大値	18630000	2988	22	0.38	0.75	0.25	0.58

表6.5.3 データ数（カテゴリカルデータ）

	組織ダミー	シート利用ダミー
0	1080	1047
1	299	332

6.5.3 回答者属性による回答傾向の比較結果（傾向スコアマッチング）

傾向スコアを用いたマッチングの結果、経験年数5年未満の群と5年以上の群にそれぞれ430ずつ、合計860のデータセットが得られた。このデータセットについての分析結果を以下に示していく。まず、表6.5.4はマッチング前後での各共変量の標準化平均差を示している。

表6.5.4 各共変量の標準化平均差

変数	マッチング前	マッチング後
	標準化平均差	標準化平均差
管轄人口	0.1748	0.0596
現在対応中の対応件数	0.26	0.0016
18歳未満在住比率	0.0584	0.0942
核世帯家族比率	0.1172	0.0901
20歳未満人口割合	0.2272	0.1419
65歳以上人口割合	0.2375	0.0778
組織ダミー	0.7063	0.4917
シート利用ダミー	0.4018	0.2513

前述の通り、2群間の共変量のバランスが取れていることを確認する際、各変数における標準化平均差は0.1以下であることが推奨されている。ここで、表6.5.4より、20歳未満人口割合、組織ダミー、シート利用ダミーについて標準化平均差は0.1を上回る結果となった。また、マッチング前後での標準化平均差を比較すると、概ねマッチング後の方が標準化平均差は小さくなっており、マッチング後の方が共変量のバランスが取れていることが分かる。しかし18歳未満在住比率についてはマッチング前に比べマッチング後の標準化平均差が大きい結果となった。

続いて、マッチング後の経験年数5年未満の群と5年以上の群について、仮想事例別に20個の必須アセスメント項目に関する回答の傾向に違いがあるかどうかの調査結果を示す。マッチングを行った後のデータセットを用いて、仮想事例別、項目別に経験年数5年未満と5年以上での回答比率を集計した。表6.5.5は仮想事例1から仮想事例3、表6.5.6は仮想事例4から仮想事例6、表6.5.7は仮想事例7から仮想事例9、表6.5.8は仮想事例9から仮想事例12における結果を示している。なお、各仮想事例のデータ数は表6.5.9に示した。

表6.5.5 経験年数5年未満と5年以上での
アセスメント項目1～20の回答比率（仮想事例1-仮想事例3）

事例1

事例2

事例3

項目 (値は小計第3位で四捨五入済み)				項目 (値は小計第3位で四捨五入済み)				項目 (値は小計第3位で四捨五入済み)						
事例1	該当	非該当	不明/情報未取得	事例2	該当	非該当	不明/情報未取得	事例3	該当	非該当	不明/情報未取得			
1	5年未満(n=58) 5年以上(n=42)	0.52 0.74	0.24 0.12	0.24 0.14	1	5年未満(n=23) 5年以上(n=24)	0.57 0.75	0.09 0.04	0.35 0.21	1	5年未満(n=57) 5年以上(n=54)	0.23 0.31	0.37 0.17	不明/情報未取得 0.40 0.52
2	事例1 5年未満(n=58) 5年以上(n=42)	0.53 0.69	0.10 0.05	0.36 0.26	2	事例2 5年未満(n=23) 5年以上(n=24)	0.91 0.96	0.00 0.00	不明/情報未取得 0.04	2	事例3 5年未満(n=57) 5年以上(n=54)	0.19 0.09	0.21 0.24	不明/情報未取得 0.60 0.67
3	事例1 5年未満(n=58) 5年以上(n=42)	0.36 0.48	0.26 0.21	不明/情報未取得 0.31	3	事例2 5年未満(n=23) 5年以上(n=24)	0.65 1.00	0.09 0.00	不明/情報未取得 0.26 0.00	3	事例3 5年未満(n=57) 5年以上(n=54)	0.47 0.43	0.12 0.07	不明/情報未取得 0.40 0.50
4	事例1 5年未満(n=58) 5年以上(n=42)	0.86 0.93	0.07 0.02	不明/情報未取得 0.05	4	事例2 5年未満(n=23) 5年以上(n=24)	0.26 0.46	0.39 0.12	不明/情報未取得 0.35 0.42	4	事例3 5年未満(n=57) 5年以上(n=54)	0.46 0.39	0.16 0.13	不明/情報未取得 0.39 0.48
5	事例1 5年未満(n=58) 5年以上(n=42)	0.45 0.43	0.16 0.17	不明/情報未取得 0.40	5	事例2 5年未満(n=23) 5年以上(n=24)	0.91 0.83	0.04 0.00	不明/情報未取得 0.04 0.17	5	事例3 5年未満(n=57) 5年以上(n=54)	0.16 0.17	0.04 0.02	不明/情報未取得 0.81
6	事例1 5年未満(n=58) 5年以上(n=42)	0.28 0.31	0.24 0.14	不明/情報未取得 0.55	6	事例2 5年未満(n=23) 5年以上(n=24)	0.57 0.79	0.22 0.08	不明/情報未取得 0.12	6	事例3 5年未満(n=57) 5年以上(n=54)	0.79 0.83	0.07 0.00	不明/情報未取得 0.14 0.17
7	事例1 5年未満(n=58) 5年以上(n=42)	0.52 0.62	0.07 0.05	不明/情報未取得 0.33	7	事例2 5年未満(n=23) 5年以上(n=24)	0.52 0.71	0.09 0.00	不明/情報未取得 0.39 0.29	7	事例3 5年未満(n=57) 5年以上(n=54)	0.19 0.33	0.07 0.09	不明/情報未取得 0.74 0.57
8	事例1 5年未満(n=58) 5年以上(n=42)	0.74 0.81	0.10 0.12	不明/情報未取得 0.07	8	事例2 5年未満(n=23) 5年以上(n=24)	0.00 0.00	0.87 0.62	不明/情報未取得 0.13 0.38	8	事例3 5年未満(n=57) 5年以上(n=54)	0.00 0.06	0.65 0.50	不明/情報未取得 0.35 0.44
9	事例1 5年未満(n=58) 5年以上(n=42)	0.88 0.90	0.00 0.07	不明/情報未取得 0.02	9	事例2 5年未満(n=23) 5年以上(n=24)	0.83 0.79	0.09 0.04	不明/情報未取得 0.09 0.17	9	事例3 5年未満(n=57) 5年以上(n=54)	0.07 0.11	0.14 0.09	不明/情報未取得 0.79 0.80
10	事例1 5年未満(n=58) 5年以上(n=42)	0.98 0.95	0.00 0.00	不明/情報未取得 0.05	10	事例2 5年未満(n=23) 5年以上(n=24)	0.09 0.12	0.26 0.17	不明/情報未取得 0.65 0.71	10	事例3 5年未満(n=57) 5年以上(n=54)	0.02 0.07	0.14 0.09	不明/情報未取得 0.84 0.83
11	事例1 5年未満(n=58) 5年以上(n=42)	0.50 0.50	0.29 0.31	不明/情報未取得 0.19	11	事例2 5年未満(n=23) 5年以上(n=24)	0.09 0.04	0.52 0.46	不明/情報未取得 0.39 0.50	11	事例3 5年未満(n=57) 5年以上(n=54)	0.00 0.00	0.61 0.59	不明/情報未取得 0.39 0.41
12	事例1 5年未満(n=58) 5年以上(n=42)	0.53 0.43	0.24 0.31	不明/情報未取得 0.26	12	事例2 5年未満(n=23) 5年以上(n=24)	0.83 0.88	0.04 0.00	不明/情報未取得 0.13 0.12	12	事例3 5年未満(n=57) 5年以上(n=54)	0.12 0.11	0.18 0.17	不明/情報未取得 0.70 0.72
13	事例1 5年未満(n=58) 5年以上(n=42)	0.52 0.36	0.12 0.21	不明/情報未取得 0.43	13	事例2 5年未満(n=23) 5年以上(n=24)	0.04 0.17	0.22 0.08	不明/情報未取得 0.74 0.75	13	事例3 5年未満(n=57) 5年以上(n=54)	0.21 0.13	0.12 0.07	不明/情報未取得 0.67 0.80
14	事例1 5年未満(n=58) 5年以上(n=42)	0.78 0.71	0.03 0.10	不明/情報未取得 0.19	14	事例2 5年未満(n=23) 5年以上(n=24)	0.04 0.08	0.48 0.38	不明/情報未取得 0.46 0.54	14	事例3 5年未満(n=57) 5年以上(n=54)	0.00 0.00	0.42 0.31	不明/情報未取得 0.58 0.69
15	事例1 5年未満(n=58) 5年以上(n=42)	0.57 0.62	0.07 0.10	不明/情報未取得 0.29	15	事例2 5年未満(n=23) 5年以上(n=24)	0.74 0.83	0.22 0.00	不明/情報未取得 0.04 0.17	15	事例3 5年未満(n=57) 5年以上(n=54)	0.07 0.04	0.77 0.69	不明/情報未取得 0.16 0.28
16	事例1 5年未満(n=58) 5年以上(n=42)	0.88 0.90	0.02 0.00	不明/情報未取得 0.10	16	事例2 5年未満(n=23) 5年以上(n=24)	0.57 0.71	0.35 0.12	不明/情報未取得 0.09 0.17	16	事例3 5年未満(n=57) 5年以上(n=54)	0.93 0.94	0.04 0.02	不明/情報未取得 0.04 0.04
17	事例1 5年未満(n=58) 5年以上(n=42)	0.00 0.00	0.76 0.83	不明/情報未取得 0.17	17	事例2 5年未満(n=23) 5年以上(n=24)	0.00 0.00	0.70 0.67	不明/情報未取得 0.30 0.33	17	事例3 5年未満(n=57) 5年以上(n=54)	0.00 0.04	0.61 0.54	不明/情報未取得 0.39 0.43
18	事例1 5年未満(n=58) 5年以上(n=42)	0.00 0.00	0.47 0.24	不明/情報未取得 0.76	18	事例2 5年未満(n=23) 5年以上(n=24)	0.00 0.00	0.65 0.29	不明/情報未取得 0.35 0.71	18	事例3 5年未満(n=57) 5年以上(n=54)	0.00 0.00	0.33 0.43	不明/情報未取得 0.67 0.57
19	事例1 5年未満(n=58) 5年以上(n=42)	0.09 0.10	0.12 0.14	不明/情報未取得 0.76	19	事例2 5年未満(n=23) 5年以上(n=24)	0.04 0.00	0.26 0.38	不明/情報未取得 0.70 0.62	19	事例3 5年未満(n=57) 5年以上(n=54)	0.35 0.57	0.12 0.04	不明/情報未取得 0.53 0.39
20	事例1 5年未満(n=58) 5年以上(n=42)	0.57 0.67	0.05 0.07	不明/情報未取得 0.38	20	事例2 5年未満(n=23) 5年以上(n=24)	0.00 0.00	0.26 0.33	不明/情報未取得 0.74 0.67	20	事例3 5年未満(n=57) 5年以上(n=54)	0.00 0.00	0.32 0.26	不明/情報未取得 0.68 0.74

表6.5.6 経験年数5年未満と5年以上での
アセスメント項目1～20の回答比率（仮想事例4-仮想事例6）

事例4

事例5

事例6

項目 (値は小數第3位で四捨五入済み)				項目 (値は小數第3位で四捨五入済み)				項目 (値は小數第3位で四捨五入済み)				
項目	該当	非該当	不明/情報未取得	項目	該当	非該当	不明/情報未取得	項目	該当	非該当	不明/情報未取得	
1	事例4			事例5				事例6				
	5年未満(n=25)	0.56	0.08	0.36	5年未満(n=44)	0.70	0.07	0.23	5年未満(n=19)	0.68	0.11	0.21
	5年以上(n=27)	0.81	0.11	0.07	5年以上(n=52)	0.81	0.06	0.13	5年以上(n=25)	0.92	0.04	0.04
2	事例4			事例5				事例6				
	5年未満(n=25)	0.36	0.16	0.48	5年未満(n=44)	0.55	0.20	0.25	5年未満(n=19)	0.26	0.47	0.26
	5年以上(n=27)	0.33	0.07	0.59	5年以上(n=52)	0.77	0.13	0.10	5年以上(n=25)	0.20	0.36	0.44
3	事例4			事例5				事例6				
	5年未満(n=25)	0.84	0.04	0.12	5年未満(n=44)	0.75	0.07	0.18	5年未満(n=19)	0.32	0.37	0.32
	5年以上(n=27)	1.00	0.00	0.00	5年以上(n=52)	0.85	0.04	0.12	5年以上(n=25)	0.56	0.04	0.40
4	事例4			事例5				事例6				
	5年未満(n=25)	0.76	0.00	0.24	5年未満(n=44)	0.07	0.16	0.77	5年未満(n=19)	0.05	0.37	0.58
	5年以上(n=27)	0.92	0.04	0.04	5年以上(n=52)	0.08	0.02	0.90	5年以上(n=25)	0.00	0.16	0.84
5	事例4			事例5				事例6				
	5年未満(n=25)	0.56	0.00	0.44	5年未満(n=44)	0.93	0.00	0.07	5年未満(n=19)	0.68	0.05	0.26
	5年以上(n=27)	0.81	0.00	0.19	5年以上(n=52)	0.96	0.00	0.04	5年以上(n=25)	0.68	0.04	0.28
6	事例4			事例5				事例6				
	5年未満(n=25)	1.00	0.00	0.00	5年未満(n=44)	0.98	0.00	0.02	5年未満(n=19)	0.89	0.00	0.11
	5年以上(n=27)	1.00	0.00	0.00	5年以上(n=52)	0.98	0.00	0.02	5年以上(n=25)	0.96	0.00	0.04
7	事例4			事例5				事例6				
	5年未満(n=25)	0.24	0.08	0.68	5年未満(n=44)	0.61	0.07	0.32	5年未満(n=19)	0.37	0.32	0.32
	5年以上(n=27)	0.19	0.07	0.74	5年以上(n=52)	0.69	0.04	0.27	5年以上(n=25)	0.68	0.08	0.24
8	事例4			事例5				事例6				
	5年未満(n=25)	0.08	0.32	0.60	5年未満(n=44)	0.00	0.39	0.61	5年未満(n=19)	0.05	0.32	0.63
	5年以上(n=27)	0.07	0.41	0.52	5年以上(n=52)	0.02	0.25	0.73	5年以上(n=25)	0.16	0.36	0.48
9	事例4			事例5				事例6				
	5年未満(n=25)	0.28	0.08	0.64	5年未満(n=44)	0.80	0.00	0.20	5年未満(n=19)	0.21	0.16	0.63
	5年以上(n=27)	0.22	0.15	0.63	5年以上(n=52)	0.81	0.00	0.19	5年以上(n=25)	0.24	0.32	0.44
10	事例4			事例5				事例6				
	5年未満(n=25)	0.04	0.16	0.80	5年未満(n=44)	0.82	0.00	0.18	5年未満(n=19)	0.26	0.16	0.58
	5年以上(n=27)	0.04	0.19	0.78	5年以上(n=52)	0.87	0.02	0.12	5年以上(n=25)	0.16	0.24	0.60
11	事例4			事例5				事例6				
	5年未満(n=25)	0.00	0.44	0.56	5年未満(n=44)	0.02	0.43	0.55	5年未満(n=19)	0.00	0.58	0.42
	5年以上(n=27)	0.00	0.67	0.33	5年以上(n=52)	0.02	0.40	0.58	5年以上(n=25)	0.00	0.52	0.48
12	事例4			事例5				事例6				
	5年未満(n=25)	0.76	0.08	0.16	5年未満(n=44)	0.89	0.02	0.09	5年未満(n=19)	0.84	0.11	0.05
	5年以上(n=27)	0.74	0.04	0.22	5年以上(n=52)	0.88	0.02	0.10	5年以上(n=25)	0.88	0.08	0.04
13	事例4			事例5				事例6				
	5年未満(n=25)	0.88	0.04	0.08	5年未満(n=44)	0.39	0.09	0.52	5年未満(n=19)	0.95	0.00	0.05
	5年以上(n=27)	0.93	0.00	0.07	5年以上(n=52)	0.50	0.02	0.48	5年以上(n=25)	1.00	0.00	0.00
14	事例4			事例5				事例6				
	5年未満(n=25)	0.04	0.20	0.76	5年未満(n=44)	0.00	0.14	0.86	5年未満(n=19)	0.11	0.16	0.74
	5年以上(n=27)	0.07	0.26	0.67	5年以上(n=52)	0.04	0.15	0.81	5年以上(n=25)	0.12	0.20	0.68
15	事例4			事例5				事例6				
	5年未満(n=25)	0.60	0.04	0.36	5年未満(n=44)	0.45	0.02	0.52	5年未満(n=19)	0.11	0.21	0.68
	5年以上(n=27)	0.74	0.00	0.26	5年以上(n=52)	0.50	0.04	0.46	5年以上(n=25)	0.20	0.32	0.48
16	事例4			事例5				事例6				
	5年未満(n=25)	0.96	0.00	0.04	5年未満(n=44)	0.43	0.16	0.41	5年未満(n=19)	0.11	0.42	0.47
	5年以上(n=27)	0.85	0.00	0.15	5年以上(n=52)	0.44	0.04	0.52	5年以上(n=25)	0.16	0.24	0.60
17	事例4			事例5				事例6				
	5年未満(n=25)	0.60	0.20	0.20	5年未満(n=44)	0.27	0.36	0.36	5年未満(n=19)	0.47	0.47	0.05
	5年以上(n=27)	0.85	0.07	0.07	5年以上(n=52)	0.40	0.35	0.25	5年以上(n=25)	0.60	0.32	0.08
18	事例4			事例5				事例6				
	5年未満(n=25)	0.28	0.16	0.56	5年未満(n=44)	0.50	0.16	0.34	5年未満(n=19)	0.05	0.37	0.58
	5年以上(n=27)	0.15	0.30	0.56	5年以上(n=52)	0.69	0.02	0.29	5年以上(n=25)	0.04	0.24	0.72
19	事例4			事例5				事例6				
	5年未満(n=25)	1.00	0.00	0.00	5年未満(n=44)	0.98	0.00	0.02	5年未満(n=19)	0.58	0.05	0.37
	5年以上(n=27)	1.00	0.00	0.00	5年以上(n=52)	0.98	0.00	0.02	5年以上(n=25)	0.68	0.08	0.24
20	事例4			事例5				事例6				
	5年未満(n=25)	0.44	0.08	0.48	5年未満(n=44)	0.23	0.34	0.43	5年未満(n=19)	0.53	0.16	0.32
	5年以上(n=27)	0.41	0.15	0.44	5年以上(n=52)	0.40	0.25	0.35	5年以上(n=25)	0.40	0.12	0.48

表6.5.7 経験年数5年未満と5年以上での
アセスメント項目1～20の回答比率（仮想事例7-仮想事例9）

事例7

事例8

事例9

項目 (値は小計第3位で四捨五入済み)				項目 (値は小計第3位で四捨五入済み)				項目 (値は小計第3位で四捨五入済み)				
項目	該当	非該当	不明/情報未取得	項目	該当	非該当	不明/情報未取得	項目	該当	非該当	不明/情報未取得	
1	事例7			事例8				事例9				
	5年未満(n=45)	0.93	0.04	0.02	5年未満(n=21)	0.62	0.19	0.19	5年未満(n=48)	0.75	0.10	0.15
	5年以上(n=50)	0.82	0.04	0.14	5年以上(n=24)	0.67	0.25	0.08	5年以上(n=45)	0.78	0.09	0.13
2	事例7			事例8				事例9				
	5年未満(n=45)	0.42	0.20	0.38	5年未満(n=21)	0.90	0.10	0.10	5年未満(n=48)	0.25	0.17	0.58
	5年以上(n=50)	0.32	0.24	0.44	5年以上(n=24)	0.75	0.04	0.21	5年以上(n=45)	0.18	0.13	0.69
3	事例7			事例8				事例9				
	5年未満(n=45)	0.96	0.02	0.02	5年未満(n=21)	0.24	0.29	0.48	5年未満(n=48)	0.77	0.04	0.19
	5年以上(n=50)	0.98	0.02	0.00	5年以上(n=24)	0.46	0.08	0.46	5年以上(n=45)	0.76	0.07	0.18
4	事例7			事例8				事例9				
	5年未満(n=45)	0.09	0.20	0.71	5年未満(n=21)	0.95	0.00	0.05	5年未満(n=48)	0.67	0.08	0.25
	5年以上(n=50)	0.02	0.28	0.70	5年以上(n=24)	1.00	0.00	0.00	5年以上(n=45)	0.64	0.09	0.27
5	事例7			事例8				事例9				
	5年未満(n=45)	0.44	0.00	0.56	5年未満(n=21)	0.48	0.24	0.29	5年未満(n=48)	0.62	0.00	0.38
	5年以上(n=50)	0.52	0.00	0.48	5年以上(n=24)	0.42	0.29	0.29	5年以上(n=45)	0.76	0.00	0.24
6	事例7			事例8				事例9				
	5年未満(n=45)	0.87	0.02	0.11	5年未満(n=21)	0.43	0.19	0.38	5年未満(n=48)	0.73	0.04	0.23
	5年以上(n=50)	0.76	0.00	0.24	5年以上(n=24)	0.42	0.12	0.46	5年以上(n=45)	0.67	0.11	0.22
7	事例7			事例8				事例9				
	5年未満(n=45)	0.33	0.04	0.62	5年未満(n=21)	0.95	0.00	0.05	5年未満(n=48)	0.50	0.02	0.48
	5年以上(n=50)	0.32	0.10	0.58	5年以上(n=24)	1.00	0.00	0.00	5年以上(n=45)	0.40	0.02	0.58
8	事例7			事例8				事例9				
	5年未満(n=45)	0.18	0.49	0.33	5年未満(n=21)	0.52	0.43	0.05	5年未満(n=48)	0.64	0.17	0.19
	5年以上(n=50)	0.24	0.38	0.38	5年以上(n=24)	0.33	0.46	0.21	5年以上(n=45)	0.02	0.76	0.22
9	事例7			事例8				事例9				
	5年未満(n=45)	0.24	0.11	0.64	5年未満(n=21)	0.57	0.05	0.38	5年未満(n=48)	0.23	0.10	0.67
	5年以上(n=50)	0.20	0.08	0.72	5年以上(n=24)	0.67	0.00	0.33	5年以上(n=45)	0.20	0.02	0.78
10	事例7			事例8				事例9				
	5年未満(n=45)	0.07	0.27	0.67	5年未満(n=21)	0.67	0.00	0.33	5年未満(n=48)	0.31	0.19	0.50
	5年以上(n=50)	0.08	0.26	0.66	5年以上(n=24)	0.71	0.00	0.29	5年以上(n=45)	0.33	0.11	0.56
11	事例7			事例8				事例9				
	5年未満(n=45)	0.09	0.53	0.38	5年未満(n=21)	0.62	0.05	0.10	5年未満(n=48)	0.29	0.23	0.48
	5年以上(n=50)	0.02	0.54	0.44	5年以上(n=24)	0.62	0.00	0.38	5年以上(n=45)	0.33	0.11	0.56
12	事例7			事例8				事例9				
	5年未満(n=45)	0.29	0.13	0.58	5年未満(n=21)	0.57	0.19	0.24	5年未満(n=48)	0.85	0.04	0.10
	5年以上(n=50)	0.32	0.14	0.54	5年以上(n=24)	0.50	0.29	0.21	5年以上(n=45)	0.78	0.07	0.16
13	事例7			事例8				事例9				
	5年未満(n=45)	0.62	0.04	0.33	5年未満(n=21)	0.71	0.05	0.24	5年未満(n=48)	0.96	0.00	0.04
	5年以上(n=50)	0.62	0.08	0.30	5年以上(n=24)	0.46	0.00	0.54	5年以上(n=45)	1.00	0.00	0.00
14	事例7			事例8				事例9				
	5年未満(n=45)	0.16	0.27	0.58	5年未満(n=21)	0.05	0.62	0.33	5年未満(n=48)	0.06	0.17	0.77
	5年以上(n=50)	0.20	0.22	0.58	5年以上(n=24)	0.00	0.62	0.38	5年以上(n=45)	0.02	0.20	0.78
15	事例7			事例8				事例9				
	5年未満(n=45)	0.98	0.00	0.02	5年未満(n=21)	0.72	0.14	0.14	5年未満(n=48)	0.40	0.12	0.48
	5年以上(n=50)	0.96	0.04	0.00	5年以上(n=24)	0.46	0.25	0.29	5年以上(n=45)	0.31	0.16	0.53
16	事例7			事例8				事例9				
	5年未満(n=45)	0.56	0.09	0.36	5年未満(n=21)	0.00	0.33	0.67	5年未満(n=48)	0.25	0.04	0.71
	5年以上(n=50)	0.46	0.08	0.46	5年以上(n=24)	0.04	0.42	0.54	5年以上(n=45)	0.27	0.02	0.71
17	事例7			事例8				事例9				
	5年未満(n=45)	0.02	0.51	0.47	5年未満(n=21)	0.05	0.90	0.05	5年未満(n=48)	0.00	0.19	0.21
	5年以上(n=50)	0.02	0.48	0.50	5年以上(n=24)	0.00	1.00	0.00	5年以上(n=45)	0.02	0.36	0.02
18	事例7			事例8				事例9				
	5年未満(n=45)	0.84	0.07	0.09	5年未満(n=21)	0.00	0.52	0.48	5年未満(n=48)	0.00	0.42	0.58
	5年以上(n=50)	0.84	0.06	0.10	5年以上(n=24)	0.00	0.42	0.58	5年以上(n=45)	0.02	0.38	0.60
19	事例7			事例8				事例9				
	5年未満(n=45)	0.67	0.07	0.27	5年未満(n=21)	0.10	0.29	0.61	5年未満(n=48)	0.69	0.04	0.27
	5年以上(n=50)	0.56	0.02	0.42	5年以上(n=24)	0.17	0.29	0.54	5年以上(n=45)	0.58	0.11	0.31
20	事例7			事例8				事例9				
	5年未満(n=45)	0.53	0.07	0.40	5年未満(n=21)	0.57	0.14	0.29	5年未満(n=48)	0.58	0.10	0.31
	5年以上(n=50)	0.50	0.02	0.48	5年以上(n=24)	0.42	0.17	0.42	5年以上(n=45)	0.58	0.16	0.27

表6.5.8 経験年数5年未満と5年以上での
アセスメント項目1～20の回答比率（仮想事例10-仮想事例12）

事例10

事例11

事例12

事例10 (値は小計第3位で四捨五入済み)				事例11 (値は小計第3位で四捨五入済み)				事例12 (値は小計第3位で四捨五入済み)						
項目	該当	非該当	不明/情報未取得	項目	該当	非該当	不明/情報未取得	項目	該当	非該当	不明/情報未取得			
1	5年未満(n=23)	0.74	0.04	0.22	1	5年未満(n=45)	0.82	0.11	0.07	1	5年未満(n=22)	0.73	0.05	0.23
	5年以上(n=24)	0.83	0.04	0.12		5年以上(n=45)	0.93	0.00	0.07		5年以上(n=18)	0.89	0.00	0.11
	不明/情報未取得					不明/情報未取得					不明/情報未取得			

表6.5.9 仮想事例別での回答者(組織)数

仮想事例番号	回答者属性別 回答者(組織)数		合計回答者(組織)数
	経験年数5年未満	経験年数5年以上	
1	58	42	100
2	23	24	47
3	57	54	111
4	25	27	52
5	44	52	96
6	19	25	44
7	45	50	95
8	21	24	45
9	48	45	93
10	23	24	47
11	45	45	90
12	22	18	40

また、それぞれの項目について、経験年数5年未満と5年以上で回答に統計的に有意な違いが存在するかを確かめるために自由度2、有意水準5%でのカイ2乗検定を行った。表6.5.10はその結果である。なお、各行はアセスメント項目、各列は仮想事例番号を示し、セル内の上側の値は χ^2 値、下側の値は p 値を示している。また、有意水準5%で有意 (p 値 <0.05) だった箇所を赤字で表示している。

表6.5.10カイ2乗検定の結果

(上の値が χ^2 値、下の値が p 値、有意水準5%で有意だった箇所を赤字で表示)

	仮想事例1	仮想事例2	仮想事例3	仮想事例4	仮想事例5	仮想事例6	仮想事例7	仮想事例8	仮想事例9	仮想事例10	仮想事例11	仮想事例12
項目1	9.20 (0.01)	6.69 (0.04)	17.77 (0.00)	3.14 (0.21)	3.07 (0.22)	22.13 (0.00)	9.86 (0.01)	0.42 (0.81)	4.0 (0.14)	2.78 (0.25)	5.0 (0.08)	5.94 (0.05)
項目2	14.14 (0.00)	1.0 (0.32)	11.28 (0.00)	7.0 (0.03)	28.57 (0.00)	9.70 (0.01)	8.29 (0.02)	3.3 (0.19)	17.52 (0.00)	1.05 (0.59)	0.12 (0.94)	3.07 (0.22)
項目3	3.22 (0.20)	5.8 (0.06)	1.38 (0.50)	9.75 (0.01)	0.86 (0.65)	9.70 (0.00)	0.73 (0.69)	8.82 (0.01)	4.60 (0.10)	0.04 (0.98)	0.73 (0.69)	17.56 (0.00)
項目4	2.14 (0.34)	6.89 (0.03)	10.42 (0.01)	5.64 (0.06)	9.78 (0.01)	9.70 (0.01)	12.61 (0.00)	0.05 (0.98)	0.89 (0.64)	3.82 (0.15)	1.53 (0.47)	5.14 (0.08)
項目5	2.33 (0.31)	3.18 (0.20)	1.65 (0.44)	2.23 (0.33)	0.72 (0.70)	22.13 (0.00)	0.35 (0.84)	4.54 (0.10)	3.85 (0.15)	0.24 (0.89)	5.90 (0.05)	0.56 (0.75)
項目6	8.68 (0.01)	14.15 (0.00)	2.57 (0.28)	0.33 (0.85)	0.31 (0.85)	1.0 (0.61)	8.11 (0.02)	0.52 (0.77)	3.7 (0.16)	0.04 (0.98)	1.11 (0.57)	0.35 (0.84)
項目7	2.40 (0.30)	1.19 (0.55)	4.86 (0.09)	0.7 (0.70)	3.42 (0.18)	7.28 (0.03)	4.36 (0.11)	0.05 (0.98)	4.17 (0.12)	1.46 (0.48)	1.66 (0.44)	1.56 (0.46)
項目8	12.32 (0.00)	1.78 (0.41)	3.20 (0.20)	1.39 (0.50)	6.90 (0.03)	1.03 (0.60)	3.12 (0.21)	4.29 (0.12)	4.87 (0.09)	5.14 (0.08)	1.47 (0.48)	9.25 (0.01)
項目9	20.68 (0.00)	1.30 (0.52)	0.76 (0.68)	1.06 (0.59)	0.21 (0.90)	5.56 (0.06)	2.18 (0.34)	1.18 (0.55)	16.0 (0.00)	2.0 (0.37)	2.13 (0.35)	0.06 (0.97)
項目10	2.5 (0.29)	3.78 (0.15)	2.81 (0.25)	0.39 (0.82)	5.59 (0.06)	5.43 (0.07)	2.45 (0.29)	0.53 (0.77)	0.8 (0.67)	4.05 (0.13)	0.93 (0.63)	4.0 (0.14)
項目11	2.84 (0.24)	1.70 (0.43)	2.76 (0.10)	1.0 (0.61)	1.35 (0.51)	0.33 (0.85)	5.47 (0.06)	5.04 (0.08)	1.23 (0.54)	4.20 (0.12)	4.49 (0.11)	0.8 (0.67)
項目12	4.5 (0.11)	0.05 (0.98)	1.52 (0.47)	1.62 (0.45)	2.58 (0.27)	18.41 (0.00)	5.24 (0.07)	1.45 (0.48)	1.50 (0.47)	3.44 (0.18)	0.03 (0.99)	0.33 (0.85)
項目13	8.94 (0.01)	3.06 (0.22)	11.47 (0.00)	2.58 (0.27)	9.78 (0.01)	0.36 (0.84)	2.43 (0.30)	9.79 (0.01)	0.2 (0.90)	7.24 (0.03)	1.23 (0.54)	0.25 (0.88)
項目14	2.88 (0.24)	0.81 (0.67)	1.73 (0.42)	0.70 (0.71)	0.84 (0.66)	1.13 (0.57)	1.35 (0.51)	1.0 (0.61)	1.47 (0.48)	2.2 (0.33)	3.31 (0.19)	1.0 (0.61)
項目15	10.37 (0.01)	2.33 (0.31)	8.72 (0.01)	3.67 (0.16)	0.91 (0.63)	2.13 (0.34)	2.33 (0.31)	12.44 (0.00)	1.88 (0.39)	2.77 (0.25)	1.5 (0.47)	7.88 (0.02)
項目16	1.68 (0.43)	3.31 (0.19)	1.02 (0.60)	2.25 (0.33)	19.68 (0.00)	4.73 (0.09)	2.38 (0.30)	0.48 (0.79)	6.46 (0.04)	4.93 (0.09)	1.34 (0.51)	6.04 (0.05)
項目17	5.14 (0.08)	0.06 (0.97)	2.73 (0.26)	6.13 (0.05)	3.56 (0.17)	0.69 (0.71)	1.04 (0.59)	3.0 (0.22)	26.0 (0.00)	1.75 (0.42)	0.5 (0.78)	0.5 (0.78)
項目18	0.13 (0.94)	0.94 (0.62)	1.16 (0.56)	3.98 (0.14)	17.84 (0.00)	1.89 (0.39)	3.16 (0.21)	0.29 (0.87)	3.15 (0.21)	0.32 (0.85)	2.50 (0.29)	0.11 (0.95)
項目19	3.38 (0.18)	3.57 (0.17)	8.27 (0.02)	0.38 (0.83)	0.84 (0.66)	2.44 (0.30)	14.62 (0.00)	2.37 (0.31)	1.83 (0.40)	2.09 (0.35)	1.0 (0.61)	3.61 (0.16)
項目20	7.73 (0.02)	0.56 (0.75)	0.10 (0.95)	1.45 (0.49)	2.64 (0.27)	1.18 (0.55)	3.04 (0.22)	3.45 (0.18)	1.51 (0.47)	3.47 (0.18)	0.28 (0.87)	2.07 (0.36)

表6.5.10より、仮想事例1においては項目1、2、6、8、9、13、5、20、仮想事例2においては項目1、4、6、仮想事例3においては項目1、2、4、13、15、19、仮想事例4においては項目2、3、17、仮想事例5においては項目2、4、8、13、16、18、仮想事例6においては項目1、2、3、4、5、7、12、仮想事例7においては項目1、2、4、6、19、仮想事例8においては項目3、13、15、仮想事例9においては項目2、9、16、17、仮想事例10においては項目13、仮想事例12においては項目3、8、15の有意差が確認された。なお、仮想事例11においてはどの項目においても有意差は確認されなかった。

また、上記の結果を踏まえて、項目別にいくつの仮想事例において有意差が確認されたかを集計した。表6.5.11では、全部で12事例存在する仮想事例のうち、カイ2乗検定において有意水準5%で有意差が見られた事例数を項目ごとに示した。

表6.5.11 経験年数5年未満と5年以上で解答に違いが見られた事例数
(アセスメント項目別)

項目名	回答の違いが見られた事例数
1. 支援/介入の困難/資源不足	5
2. 過剰なしつけ/体罰での暴力/正座等の強制/暴力のほめかし	7
3. 養育者の都合が養育より優先/生活が自己中心的	4
4. 児童の情緒的問題/対人距離/愛着課題がある	5
5. 態度から事態改善が見込まれない	1
6. 育児スキルの不足/不履行	3
7. 過去の心理的/身体的虐待歴	1
8. 養育者を過剰に支持する	3
9. 怒りや突発的事態への対処困難	2
10. 児童の育てにくさ	0
11. 養育者に対する挑発/エスカレートする行為	0
12. 支援の拒絶/回避/無関心/支援者への態度に一貫性がない	1
13. 精神科既往歴/不安定/判断力の減退/養育困難	5
14. 養育者の家庭外ストレス	0
15. 養育負担の偏り/夜間監護がない	4
16. 説明の回避/説明内容の疑念	2
17. きょうだいの係属歴	2
18. 摂食/排泄の異常/喘息やアレルギー	1
19. 不適切な養育環境	2
20. 経済不安/就労不安定	1

表6.5.11を見ると、項目1、2、4、13は経験年数5年未満と5年以上で回答に違いが見られた仮想事例数が多い。特に項目2は半数以上の事例において回答に違いが見られた。そのため、これらの項目は該当、非該当、不明の判断基準が経験年数によって異なる可能性が考えられる。一方で項目10、11、14はどの事例においても違いは見られず、経験年数5年未満と5年以上という区分のもとでは、回答に違いがあるとはいえない。

6.6 考察（傾向スコアマッチング）

上記の結果より、まずは特に多くの仮想事例で有意差が見られたアセスメント項目2について、各仮想事例における回答比率の傾向を見ながら考察を行う。項目2において有意差が見られた仮想事例は1、3、4、5、6、7、9であるが、仮想事例1、5においては経験年数5年以上の方が「該当」と回答する比率が高く、「不明/情報未取得」と回答する比率が低くなっ

ている。一方、仮想事例3、7、9は経験年数5年未満の方が「該当」と回答する比率が高く、「不明/情報未取得」と回答する比率が低くなっている。仮想事例4、6については「該当」と回答する比率は同程度で、経験年数5年未満の方が「非該当」と回答する比率が高く、「不明/情報未取得」と回答する比率が低くなっている。ここで、項目2の内容は「過剰なしつけ/体罰での暴力/正座等の強制/暴力のほめかし」と、主に暴力行為の有無について問う項目であるが、この暴力行為について、仮想事例1では『「パパ（内縁男性）が叩いた」と本児が言った』、仮想事例5では『本児が泣き止まない場合に母が叩くこともあり』と、暴力行為が発生していることが文章で明示的に示されている。一方、その他の仮想事例(3、4、6、7、9)については、暴力行為が発生していることを暗示的に示すような事例も存在するものの、どの事例においても明示的に示す文章は存在していない。この結果から、文章において暴力行為の発生が明示的に示されている場合に、経験年数が5年以上の方がより「該当」と回答し、一方で暗示的もしくは示されていない場合は、「該当」と回答する比率は同程度か、経験年数5年未満の方が高いことが分かる。また、仮想事例によってどちらの群の回答比率が高いかは異なるものの、どの仮想事例においても経験年数5年未満と5年以上で「不明/情報未取得」と回答する比率に差があることが分かる。これらの結果から、項目2においては経験年数5年以上の群の方が、より文章中に明示的に示されている事象にのみ「該当」と回答する傾向があることが示唆される。

次に、有意差の多く見られた項目である項目1、4、13に着目し、項目2と同様に回答比率について考察を行う。まず項目1に着目し、各仮想事例における回答比率を見ると、有意差が見られた事例の内、仮想事例1、2、3、6では経験年数5年以上の方が「該当」の回答比率が高い。一方、仮想事例7は逆に経験年数5年未満の方が「該当」の回答比率が高い。また、有意差は見られなかった残りの仮想事例の全てにおいても経験年数5年以上の方が「該当」の回答比率が高い。このことから、経験年数5年以上の方が「支援/介入の困難/資源不足」について、より「該当」と回答しやすい傾向があることが示唆される。

同様に、項目4に着目して各仮想事例における回答比率を見ると、有意差が見られた事例における「該当」と「非該当」の回答比率は仮想事例により差があるものの、「不明/情報未取得」については仮想事例2、3、5、6において5年以上の方が高く、残りの仮想事例7についても経験年数5年未満と5年以上ではほぼ変わらないことが分かる。ここで、項目4の内容は『児童の情緒的問題/対人距離/愛着課題がある』と、主に児童の養育者や周囲の大人に対する情緒について問う項目であるが、有意差の見られた仮想事例2、3、5、6、7については文章中に児童の情緒について明示的に書かれた文は存在しない。この結果から、項目4においても項目2と同様に、経験年数5年以上の群の方がより文章中に明示的に示されている事象にのみ「該当」と回答する傾向があることが示唆される。

また、項目13に着目し、各仮想事例における回答比率を見ると、こちらも有意差が見られた事例における「該当」と「非該当」の回答比率は仮想事例により差があることが分かる。具体的には、「不明/情報未取得」については仮想事例1、3、8、10において5年以上の方が高く、残りの仮想事例5についても経験年数5年未満と5年以上ではほぼ変わらないことが分かる。ここで、項目13の内容は『精神科既往歴/不安定/判断力の減退/養育困難』と、主に養育者の精神的な問題について問う項目であるが、有意差の見られた仮想事例のうち、1、8、10については文章中に養育者の精神状態について書かれている文章が存在する。そのため、経験年数5年未満と5年以上では、養育者が精神的問題を抱えていると職員が判断する基準が異なる可能性がある。

上記より、経験年数が5年未満と5年以上の群では「明示的」か「非明示的」に示されているかによって、該当と答えるかどうかには差がある。5年以上の場合、明示的に示されている事象にのみ「該当」と回答する傾向がある。この結果について、5年以上の場合、「虐待誤認」の恐れも持つことで、虐待と判断することを躊躇した結果や、十分な経験に基づく精査の結果によりこのような傾向が見られた可能性がある。一方、5年未満の場合、「虐待見落とし」に対する恐れから、過剰に虐待と判断している可能性がある。したがって、非明示的な事例について判断を行う場合、経験年数の異なる職員にてチーム編成し、アセスメントを行った上で判断することが望ましいと考えられる。

第7章 全国調査結果(3)：セーフティアセスメントツールの予測的妥当性

【第7章の概要】

本事業で開発するセーフティアセスメントツールを一時保護の判断に用いるためには、それが重篤であるかどうかの判断材料となりえるのか、実際の事例で確かめる必要がある。具体的には、それぞれのアセスメント項目が該当した際に、重篤な事例であるかの予測に寄与するかどうか、すなわち「予測的妥当性」を検証する必要がある。本章では、Web調査のパート3で得られた実事例アセスメント回答データを用いて、セーフティアセスメント項目の予測的妥当性について検討する。

本章では機械学習を用いてリスク予測モデルを構築したのち、そのモデルの予測性能を参照することで、項目の予測的妥当性を評価する。結果としては、セーフティアセスメント項目の回答データから、その事例が重篤事態であるか否かを60%~70%の精度で予測できることを確認した。第8章では、この精度のもとで、重篤事態の予測に貢献する項目を抽出し、セーフティアセスメントツールを作成する。

将来的には、これらの技術を現場に導入し、人による一時保護の判断を補助することで、迅速かつ正確なセーフティアセスメントの実現を目指す。

7.1 目的：重篤事態に対するセーフティアセスメントツールの予測的妥当性

重篤な虐待事態に対するセーフティアセスメントツールの構成に際して、その予測的信頼性と妥当性を評価する必要がある。本章では第4章の解析で使用された131項目のセーフティアセスメント項目（巻末資料参照）に関して、重篤事態もしくは重篤な虐待種別の予測的妥当性を評価する。評価方針としてはまず、Web調査により得られた実事例における回答パターンのデータを用いて、セーフティアセスメント項目の回答パターンから重篤事態もしくは重篤な虐待種別を予測する機械学習モデルを構築する。その後、構築された機械学習モデルの予測精度から、セーフティアセスメント項目の予測的妥当性について評価を行う。

予測的妥当性に加えて、131項目のセーフティアセスメント項目のうち、どの項目が重篤事態もしくは重篤な虐待種別の予測に貢献しているかをSHAP (SHapley Additive exPlanations)により評価した(Scott & Lee, 2017)。SHAPによる予測的貢献度の指標は8章の「セーフティアセスメントツールの項目選定」において重要な役割を果たす。

7.1.1 一般的な妥当性の評価手続き

妥当性とは、アンケート項目などに対し、測定や評価を通し、本当に測定・評価したいものを測定できているかを表した指標である。妥当性は相関の有無をもって判断されるが、どのような変数との相関を検討するかによって、どのような妥当性を検討しているかが異なる多義的な指標である。例えば、測定したいものと類似の構成概念との相関の強さを基に妥当性を判断する指標は収束的妥当性と呼ばれる。一方で、対の構成概念との逆相関の強さを基

に妥当性を判断する指標は弁別的妥当性と呼ばれる。また、将来(先)の事象との相関の強さを基に妥当性を判断する指標は予測的妥当性と呼ばれる。

7.1.2 子ども虐待対応に係るアセスメントツールの妥当性評価について

本研究では、セーフティアセスメント項目の妥当性を予測的妥当性の観点から検討する。本研究で用いるセーフティアセスメント項目は、国内外の児童虐待に関するアセスメント項目を網羅的に収集したものである。そのため、同様のアセスメント項目を用いる収束的妥当性の検討を行おうとした場合、そもそもそのアセスメント項目の妥当性が不十分である可能性があり、本研究の趣旨に合わない。弁別的妥当性についても同様である。また、これらの外的基準に生理指標などを採用して妥当性を検討することも可能であるが、事例に紐付いた生理的指標を各関係組織が有している場合はほぼ存在しないので、これも困難である。以上の理由から、本研究では、各実事例が重篤であったかどうかの判断をセーフティアセスメント項目から予測する。これにより、各セーフティアセスメント項目の妥当性についての検討と議論を行う。

7.2 方法

7.2.1 データ処理・集計方法

本章では、Web調査のパート3（3.3.3章を参照）により得られた1447件の事例アセスメントデータを使用する。これらのデータは全国の市区町村役所児童虐待対応担当部署470箇所と児童相談所129箇所の各自治体が対応した事例情報を元に構成されている。それぞれの事例において、対象となるアセスメント項目への回答に加え、その事例が重篤な事例であるか否か（一時保護（もしくは送致）を行ったか否か）の回答結果が含まれている。さらに、重篤な事例である場合は、巻末資料(調査票の設問C01_11)に示された重篤種別それぞれについての該当状況が報告されている(複数回答式)。

各自治体には最大20事例を上限として、可能な限り「一時保護（もしくは送致）を積極的に検討したが一時保護（もしくは送致）を実施しなかった事例」と「一時保護（もしくは送致）を積極的に検討した結果、一時保護（もしくは送致）が実施された事例」の両方を含めてもらうこととした。なお、Webフォームの入力においては、その実事例の通告から保護判断までを担当した職員等、またはそれを指導・監督した職員等が回答内容を作成するよう求めた。このとき、担当事例に対して131項目すべてのアセスメント項目の回答を依頼することは叶わない。そのため、各実事例に対してどのセーフティアセスメント項目を提示するかは、パート2と同様に下表の通り提示した。各回答グループに巻末資料のセーフティアセスメント項目表の1-20の共通項目を提示したのち、表7.1に記載している独自項目番号の項目を18~19項目提示した。すなわち、各回答グループが、事例基礎情報を回答した後、計40項目程度のアセスメント項目に回答したことになる。具体的には、重要項目である項目番号1-20(共通項目)については、全組織から該当状況の報告を求めた。そして、共通項目以外の項目(独自項目)については、18~19個の項目で分割構成された6つの項目群のうち、いずれか一つのグループに含まれる各項目への回答を求めた。

表7.1 各回答グループに割り当てられた評定アセスメント項目(独自項目)の対応表

	児童相談所			表 各回答グループの割り当て表および評定項目数			備考(独自項目番号)																		
	市区町村	各グループ合計	共通項目	独自項目数	合計項目数																				
回答グループ1	36	316	352	1~20	19	39	23	26	27	29	34	35	42	43	48	52	54	62	63	64	76	91	93	97	98
回答グループ2	36	317	353	1~20	19	39	21	40	41	55	56	59	65	67	70	71	79	86	88	90	101	116	117	119	123
回答グループ3	35	316	351	1~20	18	38	25	30	36	38	46	47	50	57	58	73	75	77	89	94	96	122	125	130	
回答グループ4	36	316	352	1~20	18	38	24	32	44	45	60	66	80	82	85	87	92	99	100	102	103	105	109	110	
回答グループ5	36	316	352	1~20	19	39	22	28	51	53	61	68	72	84	95	106	108	111	112	114	115	118	120	124	126
回答グループ6	35	314	350	1~20	18	38	31	33	37	39	49	69	74	78	81	83	104	107	113	121	127	128	129	131	
合計	214	1895	2110	20	111	131																			

※各グループ内の児相・市区町村比を同数になるよう調整して、項目数に関してもグループ間で同数となるように対応させている。

7.2.2 基礎集計・母集団との対応確認

まず初めにWeb調査（パート3）で集まった実事例について各情報の基礎集計を行う。具体的には、児童年齢、主たる養育者年齢、主たる虐待者種別、その他の基礎情報（過去の係属歴、DV・面前暴力、特定妊婦、一時保護・児相送致実施）、重篤な虐待者種別に関して、その度数分布または該当件数を示す。さらに、今回の予測対象である重篤種別の分布に関しては、集まったデータが母集団を十分に近似しているかを確認する。そのために、厚生労働省による先行調査である「児童相談所での児童虐待相談対応件数<速報値>（令和元年度）

（<https://www.mhlw.go.jp/content/000696156.pdf>, Last accessed 2021.03.26）」との比較を行う。

7.2.3 機械学習によるリスク予測モデリング

本調査では、セーフティアセスメント項目の回答パターンから重篤事態もしくは重篤種別を予測するモデルとしてeXtreme Gradient Boosting(XGBoost)を採用する。再掲になるが、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)。このモデルの特徴は予測性能が比較的高い点と結果の解釈性に優れている点である。

本調査では、131のセーフティアセスメント項目それぞれに該当しているか否かの二値を説明変数とした。また、その事例が重篤事態であるか否かの二値、もしくは対象とする重篤な虐待種別に該当するか否かの二値を目的変数とする。ここで、セーフティアセスメント項目の回答が非該当である場合と不明・情報未取得の場合どちらも“該当しない”という扱いをした。

XGBoostモデルの主要なハイパーパラメータ(学習条件)は、すべての予測対象において学習率(eta)を0.05に固定した。その他の主要なハイパーパラメータである(1)条件分岐の最大の深さ(max_depth)、(2)分割によるノード重み上限(min_child_weight)、(3)データ不均衡対策の重み(scale_pos_weight)についてはグリッドサーチ(Grid Search)を用いて、後述する予測精度の一つであるPR曲線下面積(AUCPR)が最大化する組み合わせを選んだ。ブースティングのラウンド数(num_round)についてはGrid Searchにおける一回一回の試行において、交差検証法(6-fold cross validation)により最適な値を決定した。その他のハイパー

パラメータは、R言語におけるxgboostパッケージの標準設定値を利用した(Chen, T. & Guestrin, C., 2016)。

機械学習をトレーニングする際は、手元にあるデータを(1)学習データ(training)、(2)ハイパーパラメータ調整用のデータ(validation)、(3)性能評価用のデータ(test)にそれぞれ75%、15%、10%の割合で無作為に分割して作成した。このとき、重篤事態に該当する事例が学習データに少数しか含まれない場合(クラス不均衡)、解析結果に偏りが生じる恐れがある。これによる予測精度の低下を緩和するため、モデルを学習する際はクラスごとの事例数に応じた重み(scale_pos_weight)を誤差関数に付与した。

なお、本調査ではL1正則化付き回帰(LASSO)も実施したが、最終的な予測精度はXGBoostが上回ったため、XGBoostの結果を採用した。

7.2.4 予測的妥当性の評価指標

重篤事態に対する予測性能を評価する指標として、第4章でも用いたPR曲線下面積(Area Under the Precision-Recall Curve: AUCPR)を採用する。AUCPRは、横軸に再現率(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。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。クラス不均衡データでも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性能評価指標として利用が推奨されている(Sofaer et al., 2019)。

7.2.5 予測的貢献度の評価指標

各セーフティアセスメント項目が重篤事態もしくは重篤種別の予測にどの程度貢献しているかを示す評価指標として、SHAP (SHapley Additive exPlanations)を採用する(Scott & Lee, 2017)。本章でのSHAPは、重篤な事態の予測に貢献した場合は正の値を取り、重篤でない事態の予測に貢献した場合は負の値を取る数量指標である。その貢献度が大きいほど絶対値も大きくなる。

なお、予測的貢献度は本章の目的から逸脱するが、第8章の目的である項目選定において重要な役割を果たす。機械学習による解析結果を一つの章にまとめるため、便宜上、本章で予測的貢献度の結果について報告する。

7.3 結果 I : 基礎集計

最初に回収率・欠損・除外データについて述べたのち、いくつかの基礎情報に関する集計を列挙していく。

7.3.1 回収率・欠損・除外データ

Web調査のパート3において、回答者区分ごとに集計をとった結果、市区町村からは733件、児童相談所からは714件の実事例データを得た。集まった回答データにおいて、重篤な虐待種別の回答が欠損値の場合は非該当であるとして処理している。一方で、セーフティアセスメント項目の回答が欠損値の場合は不明／情報未取得として扱う。なお、他の章と同様に、回答をいただいた中の4自治体に回答不備があったため、これらの回答データは本研究の分析対象から除外した。

7.3.2 児童年齢の基礎集計

Web調査（パート3）で集まった実事例に対して、児童年齢の度数分布を図7.1に示した。6歳以上10歳未満(9歳以下)の小学校低学年に該当する児童が最も多く、全体の約24.1%を占めた(350件 / 1447件)。市区町村13件、児童相談所6件、合計19件の児童年齢欠損データが含まれた(約1.3%)。

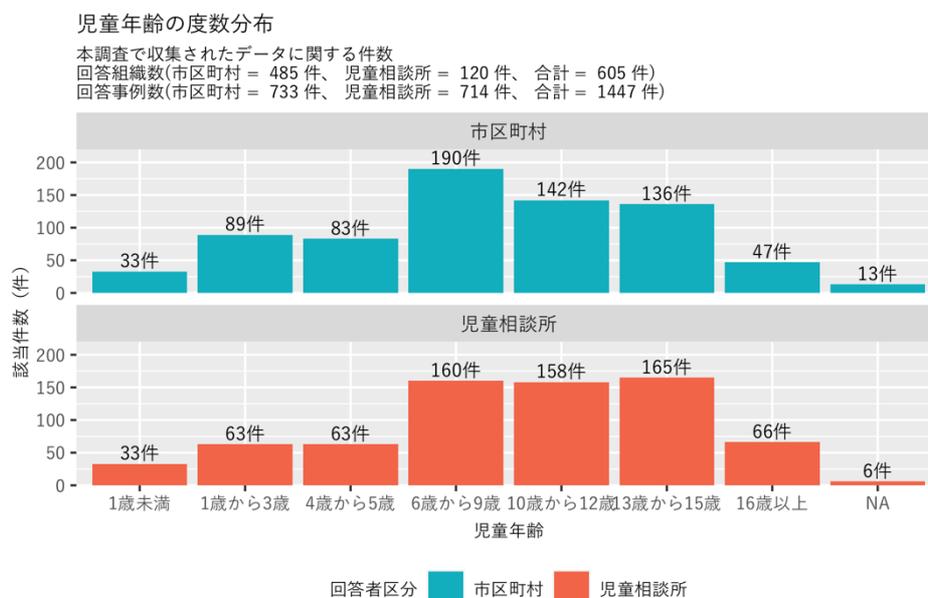


図7.1 Web調査（パート3）で集まった実事例における児童年齢の度数分布

7.3.3 主たる養育者年齢の基礎集計

主たる養育者については、35歳以上40歳未満までの年齢区分が最も多く(520件)、全体の約35.9%を占めた(図7.2)。また、市区町村14件、児童相談所5件、合計19件の欠損データが含まれた(約1.3%)。

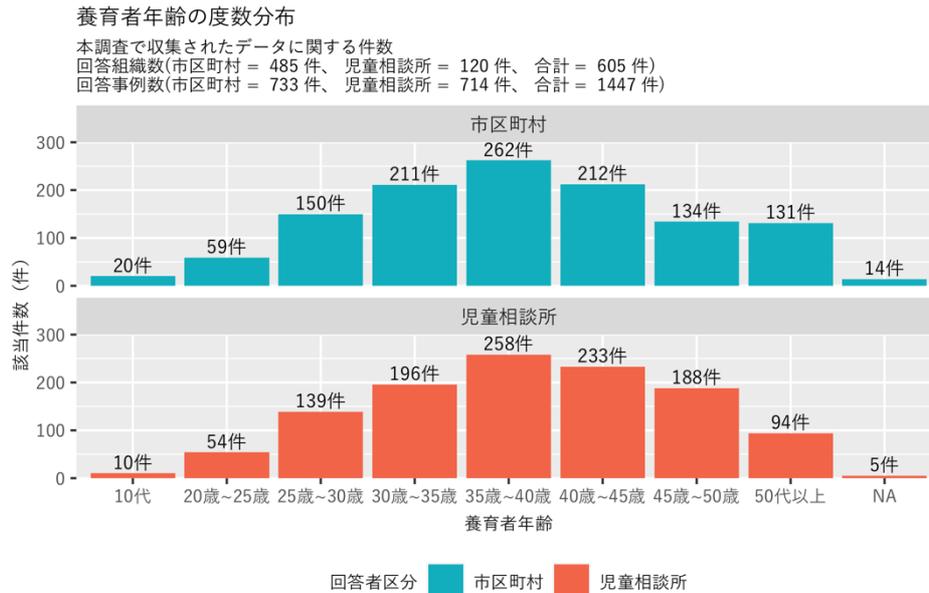


図7.2 Web調査（パート3）で集まった実事例における主たる養育者年齢の度数分布

7.3.4 主たる虐待者種別の基礎集計

主たる虐待者種別では、市区町村および児童相談所で「実母」に該当する件数が最も多く、全体の約53.9%を占めた(781件 / 1447件)。なお、(その他に含まれる内容を除き)最も少なかったのは実母以外の母親であり、全体の約2.0%であった(29件 / 1447件)。最後に欠損データは市区町村および児童相談所合わせて18件あり、全体の約1.2%であった(18件 / 1447件)。図7.3に結果を示す。

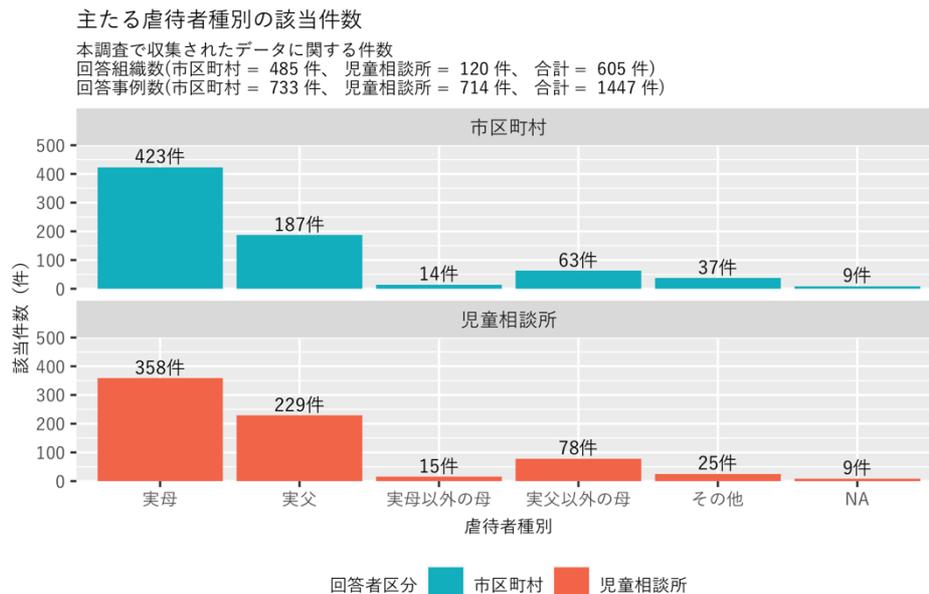


図7.3 Web調査（パート3）で集まった実事例における主たる養育者年齢の度数分布

7.3.5 その他の基礎情報項目の基礎集計

その他基礎情報項目として、虐待による過去の係属歴、DV・面前暴力、特定妊婦、一時保護の実施（または市区町村の場合、児童相談所への送致に関する項目）を取得した。それぞれの該当件数を図7.4に示す。最も該当報告件数の少ない特定妊婦において、その該当率は約12.4%となった(179件 / 1447件)。また、一時保護もしくは市区町村が児相送致を実施した事例の総数は1126件となり、全体の78.5%(1126件 / 1447件)を占めた。

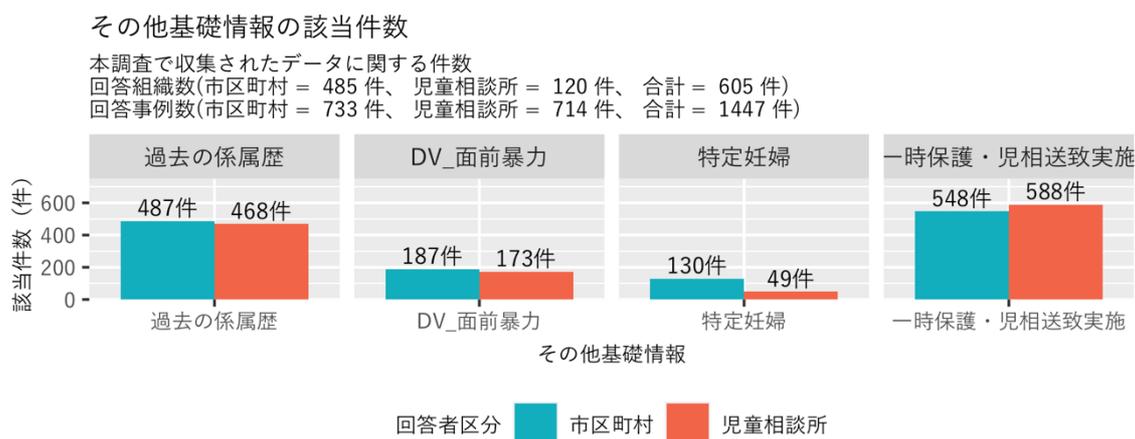


図7.4 Web調査（パート3）で集まった実事例におけるその他基礎情報の該当件数

7.3.6 重篤な虐待種別の基礎集計

重篤な虐待種別ごとの該当件数を図7.5に示す。その他の重篤項目の報告事例数が最も多く全体の44.7%（647件/1447件）を占め、重篤な性的虐待が最も少なく全体の13.2%（191件/1447件）となった。なお、複数の虐待種別が併存している事例も存在することには留意されたい。本章では「重篤な事例（重篤な虐待種別に一つでも当てはまる事例）か否か」を予測対象とする解析と、それぞれの虐待種別において「その虐待種別に該当するか否か」を予測対象とする2つの解析を実施した。

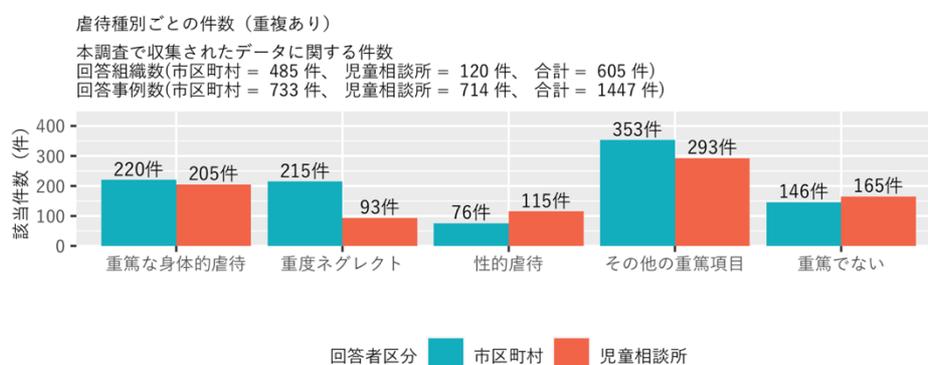


図7.5 Web調査（パート3）で集まった実事例における虐待種別ごとの該当件数

7.3.7 虐待種別に関する母集団との対応確認

厚生労働省による全国調査である「児童相談所での児童虐待相談対応件数<速報値>（令和元年度）」が虐待種別分布の母集団をよく近似していると仮定したうえで、その分布と本調査での重篤な虐待種別の分布を比較した。比較に際して、2点注意しておくべき点がある。一つは、厚生労働省の全国調査では、虐待の重篤度にかかわらず集計を行っているのに対し、本調査では重篤な事例を対象を絞っている点である。もう一つは、厚生労働省の調査では「心理的虐待」を調査対象の一つにしているが、本調査ではその代わりに「その他の重篤項目」を調査対象としている点である。こうすることで、重篤な心理的虐待を含めたより広域な重篤項目を収集するようにしている。よって、厚生労働省の全国調査における「心理的虐待」の項は「その他の重篤項目」に置き換えて比較を行った。図7.6にそれぞれの調査における虐待種別の割合を示した。性的虐待とその他の重篤項目について10ポイント以上の差が見られ、それ以外の項目では本調査における件数の割合がやや大きく出ている。これは先述したように、本調査が重篤な虐待を対象としており、比較的軽微な事態になりにくい心理的虐待が重篤でない事例に含まれていることに依ると推測できる。これを考慮すると、両者の虐待種別の分布はおおよそ対応関係があると考えられる。

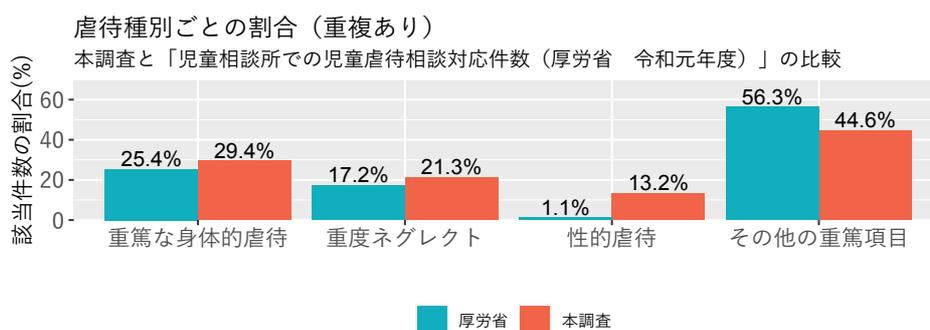


図7.6 虐待種別ごとの割合 厚生労働省による全国調査との比較

7.3.8 重篤な虐待種別の併存

ここでは、重篤な虐待種別の併存について基礎集計を行う。まず、図7.7に虐待種別が複数併存した事例の報告件数をその併存数ごとに示した。参考のため重篤でない事例（併存数が0）と、虐待種別が併存していない事例（併存数が1）の報告件数も同時に示した。併存数が大きくなるごとにその件数が減っていることが見てとれる。

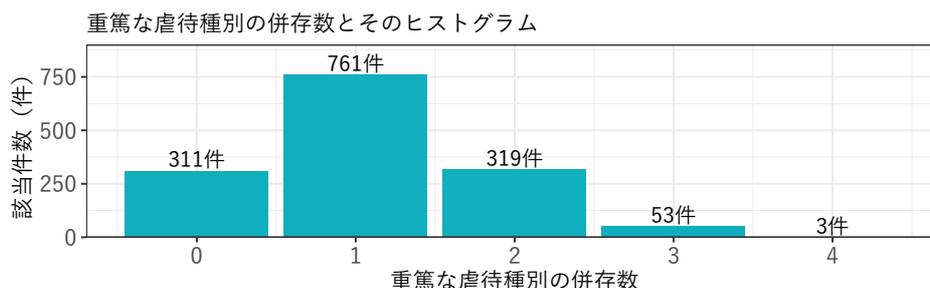


図7.7 併存数別の該当件数

次に、2つの虐待種別が併存する事例において、どの組み合わせの虐待種別が報告されたかを見る。図7.8に組み合わせごとの報告件数をクロス集計で示した。ここから、「その他の重篤項目」が他の虐待種別と併存する頻度が高いことが見てとれる。これは「その他の重篤項目」が広義な虐待を含んでいるためと推測できる。中でも、「重篤な身体的虐待」と「その他の重篤項目」の併存頻度が最も高く143件の報告があった。

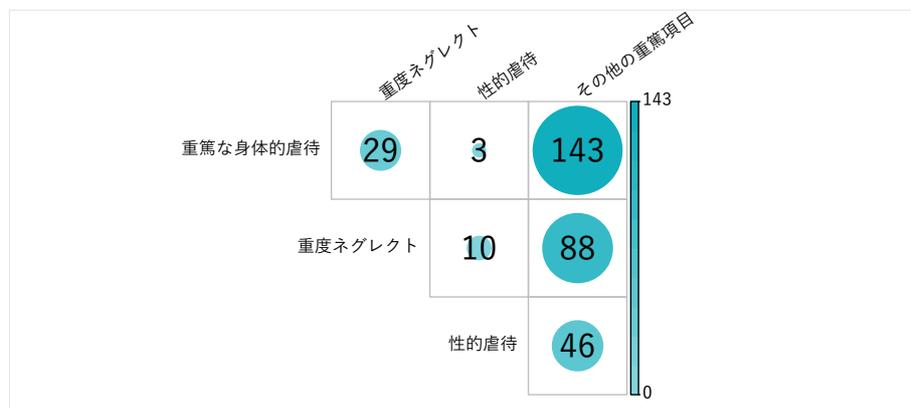


図7.8 2つの虐待種別が併存する事例における組み合わせとその件数

また、3つの虐待種別が併存した実事例において、その組み合わせごとの該当件数を表7.2にまとめた。サンプルサイズが十分とは言えないが、性的虐待以外の3つの虐待種別が併存するパターンが最も多く31件の報告があった。性的虐待は他の虐待種別と併存しにくいとの推測がたてられるが、そもそも性的虐待に該当する事例の報告が少ないため、確かではない。

表7.2 3つの虐待種別が併存した実事例の組み合わせとその件数

併存した3つの虐待種別				件数 (件)
	重度ネグレクト	性的虐待	その他の重篤項目	14
重篤な身体的虐待		性的虐待	その他の重篤項目	8
重篤な身体的虐待	重度ネグレクト		その他の重篤項目	31
重篤な身体的虐待	重度ネグレクト	性的虐待		0

7.4 結果Ⅱ：機械学習による解析

7.4.1 リスク予測モデリングの予測性能（重篤事態の予測）

7.2.3項で述べたように、Grid Searchと6-fold 交差検証を用いてハイパーパラメータを調整したうえで、セーフティアセスメント項目の回答パターンから重篤な虐待事例を予測するモデル（XGBoost）を構築した。調整後のハイパーパラメータは $\text{max_depth}=5$ 、 $\text{min_child_weight}=4$ 、 $\text{scale_pos_weight}=0.273$ となった。

検証用データに対する予測モデルのPR曲線（Precision-Recall Curve）を図7.9に示す。7.2.4項で紹介したように、PR曲線下面積は0から1の値を取り、1に近いほど予測能力が高いと言えるが、今回、その値は0.876となった。また、判別閾値を0.5とした際、「実際に重篤な事例のうち機械学習が重篤な事例であると正しく予測できた割合」を示すRecallは0.632、「機械学習が重篤であると予測した事例のうち、実際に重篤な事例であった割合」を示すPrecisionは0.857となった。最後に「与えられた重篤な事例とそうでない事例（すべての性能評価用データ）のうち、機械学習がそれらを正しく予測した割合」を示すAccuracyは0.630となった。クラス分類問題で頻繁に使われる予測評価指標のいくつかを表7.3にまとめる。

XGBoostによる重篤な虐待事例の予測精度（PR曲線）

曲線下面積 = 0.876
閾値0.5とした時の予測精度： Accuracy = 0.630, Precision = 0.857, Recall = 0.632

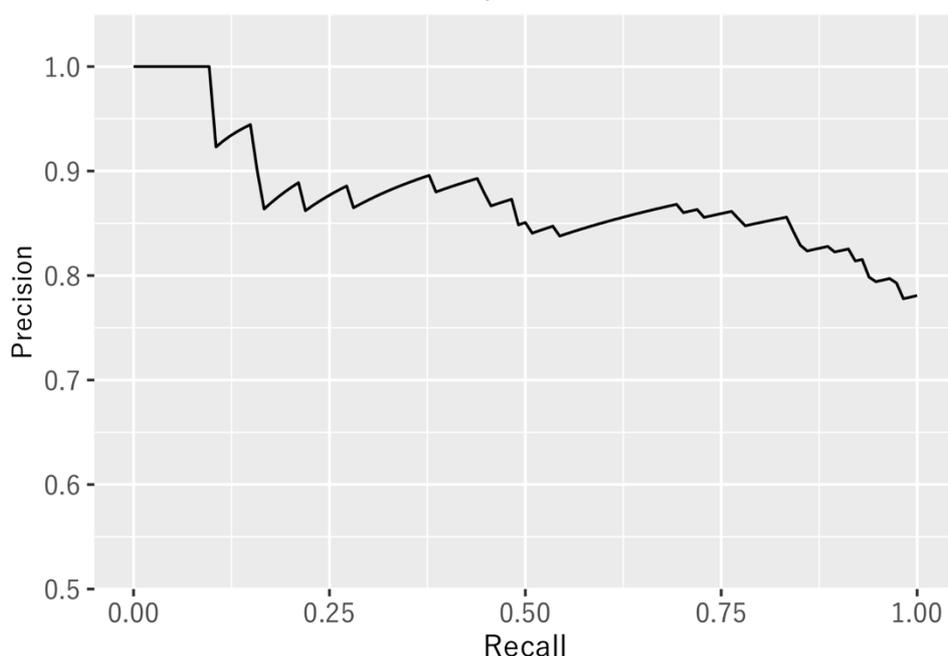


図7.9 XGBoostによる重篤な虐待事例の予測精度（PR曲線）

表7.3 判別閾値を0.5とした際の予測精度の評価指標（重篤事態）

閾値0.5とした時の予測精度						
予測対象	AUCPR	Accuracy	Precision	Recall	Specificity	Neg.Pred.Value
重篤な事態	0.876	0.63	0.857	0.632	0.625	0.323

Precisionの値がRecallやAccuracyの値に比べて大きく出ているのは、今回集まったデータにおいて、重篤でない事例のサンプルサイズが全体の21.7%と少ないことに起因する。サンプルサイズに起因する評価指標の偏りは、真のラベル（実際の事例が重篤であるか否か）と機械学習の予測を軸にとったクロス集計（混同行列）を見ることでも確認できる。ここでは、性能評価用のデータにおけるモデルの予測を用いて混同行列を評価した（図7.10）。サンプルサイズの偏りは「機械学習が重篤でないと予測した事例のうち、実際に重篤でない事例の割合（Negative Predictive Value）」が0.322と低く出ていることにも影響している。一方で「実際に重篤でない事例を機械学習が正しく重篤でないと予測できた割合（Specificity）」は0.625となり、RecallやAccuracyと同程度の値が得られている。このことから、今回構築したモデルは、60%~70%ほどの正答率であると言える。しかし実用上は、実際に重篤な事例のうち42件(36.8%)の予測に誤りがある点は問題である。それぞれのクラスの予測精度を均衡に保ちつつ、全体の予測精度を向上させるには、今後データを蓄積しサンプルサイズを増やすことが推奨される。

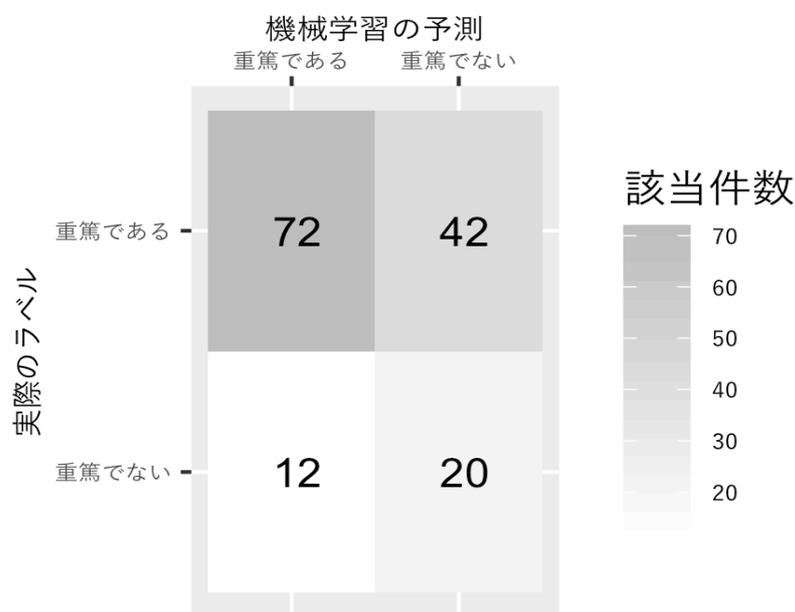


図7.10 XGboostの予測に対する混同行列

7.4.2 リスク予測モデリングの予測性能（重篤な虐待種別の予測）

7.4.1項では、重篤事態の予測を対象としたXGBoostモデルの結果を報告した。本項では同様に、それぞれの重篤な虐待種別を予測対象としたXGBoostモデルの結果を報告する。例えば「重篤な身体的虐待」を予測対象とした時は、それに該当するか否かが目的変数であり、131のセーフティアセスメント項目の回答パターンが説明変数となる。表7.4にそれぞれの予測対象における予測性能指標の結果を示す。重篤事態の予測ではそれに該当する事例が学習データの多数を占めていたが、虐待種別の予測ではそれに該当する事例が少数となっている（7.3.6項 図7.5を参照）。そのため、前項で示した重篤事態の予測性能指標（表7.3）と本項で示した表7.4は単純な比較が出来ないことに留意されたい。

学習データに含まれる該当事例数の偏り（クラス不均衡）がとくに顕著な「重度ネグレクト」と「性的虐待」ではAUCPRの値が低い。また、Precisionの値も低くなっていることから、クラス不均衡の影響が予測性能の低下を引き起こしている可能性が指摘される。この問題を解決するには、今後データを蓄積しサンプルサイズを増やすことが推奨される。

表7.4 判別閾値を0.5とした際の予測精度の評価指標（重篤な虐待種別ごと）

閾値0.5とした時の予測精度						
予測対象	AUCPR	Accuracy	Precision	Recall	Specificity	Neg.Pred.Value
重篤な身体的虐待	0.690	0.685	0.478	0.767	0.650	0.870
重度ネグレクト	0.440	0.703	0.385	0.645	0.719	0.882
性的虐待	0.343	0.699	0.260	0.650	0.706	0.927
その他の重篤項目	0.611	0.637	0.607	0.523	0.728	0.656

7.4.3 セーフティアセスメント項目の予測的貢献度（重篤事態の予測）

以上の項ではセーフティアセスメント項目の回答パターンから重篤な虐待事例を予測するモデルを構築した。本項では、どのセーフティアセスメント項目がより予測に貢献しているかを評価する。これらは、第8章で行うセーフティアセスメントツールの項目選定において参照可能な指標となる。ここでは、7.2.5項で述べたように、SHAPを予測的貢献度の指標とする。SHAPはサンプルごとの各セーフティアセスメント項目について算出される。また、セーフティアセスメント項目ごとにSHAPの絶対値をサンプルについて平均したものはGlobal SHAPと呼ばれる。この値が大きいセーフティアセスメント項目は予測に与える影響がより大きいと言える。

すでにハイパーパラメータの調整は済んでいるので、得られたデータのすべて（1447件）を用いてXGBoostの学習を行い、そのSHAPを評価した。表7.5に、Global SHAPが大きいセーフティアセスメント項目を上から順に35項目まで示した。重要項目である項目番号1～20が、比較的上位にあることが分かる。しかし、重要項目はその他の通常項目に比べて欠損値が少ないことには留意されたい。これは調査設計上、重要項目に関してはすべての対象組

織に回答をもらっている一方、その他の通常項目に関してはランダムに選ばれた特定の組織にのみ回答をもらっていることに依る。それぞれの項目の詳細については巻末資料を参照されたい。

なお、データを分割する際のランダムネスや、XGBoostの構築における初期値のランダムネスはリスク予測モデリングの予測精度に大きな影響を与えない。しかし、Global SHAPの値が比較的低い項目（貢献度の低い項目）に関しては、そのGlobal SHAPがランダムネスに影響を受けやすい（分散が大きい）。そのため、その点については結果の解釈の際に留意する必要がある。

表7.5 重篤事態の予測におけるGlobal SHAP 上位35項目

順位	項目番号	厚労省区分	統合項目名	Global.SHAP
1	1	社会関係	支援/介入の困難/資源不足	0.214
2	16	養育者の様子	説明の回避/説明内容の疑念	0.197
3	4	児童の様子	児童の情緒的問題/対人距離/愛着課題がある	0.161
4	5	養育者の様子	態度から事態改善が見込まれない	0.136
5	2	養育者の様子	過剰なしつけ/体罰での暴力/正座等の強制/暴力のほめかし	0.114
6	15	養育環境/生活状況	養育負担の偏り/夜間監視がない	0.098
7	19	養育環境/生活状況	不適切な養育環境	0.083
8	11	児童の様子	養育者に対する挑発/エスカレートする行為	0.080
9	20	家族(きょうだい)情報/世帯情報	経済不安/就労不安定	0.080
10	13	養育者の様子	精神科既往歴/不安定/判断力の減退/養育困難	0.068
11	68	児童の様子	児童の帰宅不安/恐怖	0.043
12	108	その他	今までに経験したことのない事例	0.042
13	43	重篤項目	児童が帰宅を嫌がる・拒否する	0.040
14	48	重篤項目	児童の身体に打撲痕や内出血などの外傷がある	0.040
15	37	養育者の様子	通告による傷つき/プレッシャー	0.038
16	9	養育者の様子	怒りや突発的事態への対処困難	0.037
17	18	児童の様子	摂食/排泄の異常/喘息やアレルギー	0.036
18	3	養育者の様子	養育者の都合が養育より優先/生活が自己中心的	0.032
19	98	重篤項目	児童自身が保護・救済を求めている	0.031
20	12	養育者の様子	支援の拒絶/回避/無関心/支援者への態度に一貫性がない	0.030
21	32	児童の様子	学業上での課題を抱える	0.026
22	29	養育者の様子	虐待の黙認/擁護/認識欠如	0.024
23	7	児童の様子	過去の心理的/身体的虐待歴	0.024
24	59	養育者の様子	支援者への攻撃性	0.023
25	21	児童の様子	噛み傷/わずかな傷/説明されない傷	0.022
26	10	養育者の様子	児童の育てにくさ	0.021
27	6	養育者の様子	育児スキルの不足/不履行	0.021
28	34	養育者の様子	人前での暴言暴力/泣いてもあやさない様子	0.018
29	26	家族(きょうだい)情報/世帯情報	家族構成/同居人変化	0.018
30	8	児童の様子	養育者を過剰に支持する	0.017
31	87	妊娠/出産	若年出産	0.017
32	17	家族(きょうだい)情報/世帯情報	きょうだいの係属歴	0.014
33	14	養育者の様子	養育者の家庭外ストレス	0.013
34	24	家族(きょうだい)情報/世帯情報	夫婦間葛藤/対立/話し合い困難/立場が対等ではない	0.007
35	44	養育者の様子	物質/行為依存/発達障害の診断/疑い	0.006

$$\Phi_j = \frac{1}{D} \sum_{i=1}^D x_j^{(i)} \phi_j^{(i)}$$

(式7.1)

ここで、 D はサンプルサイズ、 $x_j^{(i)} \in \{-1, +1\}$ は i 番目のデータ点における j 番目のセーフティアセスメント項目に該当する (+1) か否か (-1) を表す2値変数、 $\phi_j^{(i)} \in \mathbb{R}$ はそれに対応するSHAPである。式7.1から分かるように、データの中で $x_j^{(i)}$ と $\phi_j^{(i)}$ が同符号である頻度が高いと、Signed Global SHAPは正に大きな値をとる（例えば、項目番号1）。逆に $x_j^{(i)}$ と $\phi_j^{(i)}$ が異符号である頻度が高いと、負に大きな値をとる（例えば、項目番号19）。これは $x_j^{(i)}$ と $\phi_j^{(i)}$ の相関関係を評価していると言い換えることもできる。ここで、単にSHAPの絶対値平均をとったGlobal SHAPは「予測に与える影響」という情報を持っているのに対し、Signed Global SHAPは「予測に与える正負の影響」という情報を持っており、緩衝項目になり得る傾向を持つセーフティアセスメント項目を検出することができる。

Global SHAPの値が大きい順に上位35項目に対してSigned Global SHAPをプロットしたのが図7.12である。これを見ると、上位35項目のうち12項目でSigned Global SHAPが負となったことが分かる(図7.12の中でも、Signed Global SHAPが負となった上位10項目を表7.6に示した)。これらの項目はこの予測モデルにおける緩衝項目、つまりその項目に該当した場合に重篤である確率が下がる項目である。しかし、今回のセーフティアセスメント項目はそもそも、アセスメント項目だけを含んでいると想定している。例えば、Signed Global SHAPが負の値になる「項目番号19：不適切な養育環境」について言えば、不適切な養育環境のほうが、適切な養育環境に比べて、重篤な虐待事例が起こりやすい環境である。それにもかかわらず、この項目が予測モデルにおいては重篤である確率を下げる項目になっている原因としては、ハイコンテクストな要因が関わっていると考えられる。先の項目番号19の例で言えば、「不適切な養育環境であることを開示する養育者は支援者との信頼関係が強く、支援者と密な連携を取ることで重篤な事態を避けやすい」などのストーリーが想定できる。もちろん、これが唯一のストーリーではないが、こういった要因が積み重なり、緩衝項目になっている可能性がある。また、別の要因としては、事例のサンプルサイズ不足によるノイズやバイアスの混入が考えられる。これは特に、Global SHAPの値が低い項目に影響を与えやすく、今回集まったデータにおいて、「たまたま」緩衝項目として評価された可能性もある。

機械学習による予測モデルにおいて、各緩衝項目は、その該当を以て重篤事態の発生・併存を予測することに貢献しない。しかし以下のようなシナリオが各項目から想定されるため、現場で使用する際に、上記の項目に複合してチェックが入れられている場合には注意する必要がある。

例えば、項目番号19の「不適切な養育環境」に該当している場合、身体的、心理的、性的虐待、ネグレクトのいずれかが生じている家庭環境、もしくは、身体的、心理的、性的虐待、ネグレクトにエスカレートして繋がる行為や状態が慢性的に生じている家庭環境である。そのような家庭環境に加えて、項目番号32のような学業上での課題を抱える項目に該当している場合、学校において児童虐待の悪影響が表出されている可能性がある。例えば、子どもの学習意欲が減退したり、学業成績が悪化したりするシナリオがありえる。したがって、上記の複数の項目に複合してチェックが入れられている場合には、注意して対応する必要がある。

ここで、本節のSHAPを用いた分析においては限界が存在する。SHAPはあくまで、予測モデルをブラックボックスとして扱い、その入力値と出力値だけに着目して予測的貢献度を評価する手法である。つまりSHAPは貢献度の評価基準として確立された手法ではあるが、一方でなぜそのような予測になったかを説明することはできない。さらに、個々のサンプルに対するSHAPはある一つのセーフティアセスメント項目の貢献度を厳密に測ることができるが、その統計的代表的値であるGlobal SHAPとSigned Global SHAPに関しては、その限りでない。また、本章で示したSHAPは予測精度(Accuracy)が6~7割の予測モデルをもとに算出されていることにも注意されたい。

7.4.4 セーフティアセスメント項目の予測的貢献度（重篤な虐待種別ごとの予測）

本項では前項に引き続き、重篤な虐待種別それぞれの予測における、各セーフティアセスメント項目の予測的貢献度について示す。前項で示した重篤事態の予測で貢献度の高かった上位35項目はセーフティアセスメントツールの作成において優先度が高いとした上で、各虐待種別の予測に特有でありかつ重要な項目を抽出する。これを見るため、虐待種別それぞれについて各項目のGlobal SHAPが高い順に順位付けしたのから、前項で示した重篤事態の予測で貢献度の高かった上位35項目を除き、その上位5項目を表7.7に示した。これらは、各虐待種別の予測の際に特に注目すべき重要項目である可能性がある。

表7.7 各重篤種別の予測におけるGlobal SHAP上位5項目
(重篤事態の予測における上位35項目を除いて)

XGBoostにおけるGlobal SHAP上位項目 (重篤な身体的虐待)					
重篤事態の予測における上位35項目を除いた上位5項目					
順位	項目番号	厚労省区分	統合項目名	Global.SHAP	Signed.Global.SHAP
12	91	重篤項目	養育者が児童の身体の複数箇所を殴打している	0.065	0.064
13	69	重篤項目	頭部の瘡や抜毛(後頭部まで確認)、上まぶた・顔面に点状の出血や痣、目の血走りなど、児童の頭部・顔面に外傷がある	0.064	0.064
17	121	児童の様子	睡眠の問題/悪夢の報告	0.036	-0.036
18	95	児童の様子	学校/園での心身不調の訴え	0.036	-0.036
20	92	重篤項目	養育者に、道具を用いた体罰・暴力行為がある(軽度の外傷や跡がない場合)	0.024	0.024

XGBoostにおけるGlobal SHAP上位項目 (重度ネグレクト)					
重篤事態の予測における上位35項目を除いた上位5項目					
順位	項目番号	厚労省区分	統合項目名	Global.SHAP	Signed.Global.SHAP
11	27	重篤項目	虐待者が、虐待行為を正当化している	0.031	0.031
12	53	養育者の様子	調査協力への無理解/非協力	0.029	0.029
13	28	その他	複数種別の虐待が併発/混合	0.029	0.029
15	63	妊娠/出産	高齢出産	0.021	0.021
17	74	児童の様子	児童の発達障害(疑い含む)	0.019	-0.019

XGBoostにおけるGlobal SHAP上位項目 (性的虐待)					
重篤事態の予測における上位35項目を除いた上位5項目					
順位	項目番号	厚労省区分	統合項目名	Global.SHAP	Signed.Global.SHAP
8	35	養育者の様子	児童への絶え間ない叱責/非難/拒絶/無視	0.094	-0.094
13	76	重篤項目	虐待行為の可能性が高いと判断されるにもかかわらず、養育者が虐待を否定する・認めない	0.079	0.079
15	23	養育者の様子	うつ状態/無気力または妄想幻覚/躁状態がある	0.065	-0.065
17	55	養育者の様子	関係機関を非難/脅迫/支援の被害的受け取り	0.050	0.050
18	60	家族(きょうだい)情報/世帯情報	ひとり親家庭/夫婦間年齢差	0.049	-0.049

XGBoostにおけるGlobal SHAP上位項目 (その他の重篤項目)					
重篤事態の予測における上位35項目を除いた上位5項目					
順位	項目番号	厚労省区分	統合項目名	Global.SHAP	Signed.Global.SHAP
20	39	社会関係	近隣トラブル/養育者社会関係	0.030	0.030
24	31	養育者の様子	養育プレッシャー/育児不安/児童の障害等受容困難	0.025	0.025
26	75	重篤項目	養育者が、現状解決されていない養育上の課題に対する解決方法を求めている	0.022	0.022
27	51	養育者の様子	DVの疑い/過去のDV相談歴	0.019	0.019
28	42	重篤項目	養育者や児童が、別の養育者に対して「殺されるかもしれない」「何をするかわからない」等の確信めいた不安や恐れを訴える	0.019	0.019

有識者会議においてこれらの項目を「各虐待種別の予測に有用な可能性がある」として示したところ、「現場の感覚とは合致しない」とのご指摘をいただいた。原因としては、第一に各虐待種別の予測に貢献する項目は、重篤事態の予測に貢献するものとその大多数が重複していることが挙げられる。第二に、7.4.2項でも述べたようにサンプルサイズの不足（クラス不均衡）のために、モデルの予測精度が高くないことが原因として考えられる。ここで、現状の予測精度が低いモデルのもとでの予測的貢献度は良い指標にならない可能性がある。そのため、第8章の項目選定において、各虐待種別に関して任意入力部分は、削除することとした。なお、予測対象別にGlobal SHAPの高かった上位35項目を列挙した表は、巻末資料に掲載している。適宜参照されたい。

7.4.5 統合項目を使ったリスク予測モデリング（重篤事態の予測）

7.4.1項では重篤事態のリスク予測モデリング（XGBoost）に対して、その予測的妥当性を評価した。また、7.4.3項では構築されたXGBoostを用いてGlobal SHAPを算出し、どのセーフティアセスメント項目が重篤事態の予測に貢献しているか（予測的貢献度）を検証した。本項ではさらに、第8章の目的であるセーフティアセスメントツールの作成を見据えて、予測的貢献度が上位の項目のいくつかを統合し項目数の削減ができるかを検討する。

統合する項目は、有識者が表7.5に示した35項目の中から「類似した内容を質問している」と判断したペアを選択した。選択された項目を統合することで機械学習に inputs する項目数は減少するが、その際、予測性能に影響がでるかを検証する。

上記で述べた通り有識者が統合できる項目を検討した結果、Global SHAP上位35項目のうち項目番号3と5、項目番号43と68はそれぞれ類似の質問内容であることから統合することとなった。ここで、二つの項目を統合した統合項目は、どちらか一方の項目に該当した場合に「該当」、どちらも非該当もしくは不明の場合に「非該当もしくは不明」であるとした（論理和）。この統合を上記2つのペアに施した上で、7.4.1項と同様の手順でリスク予測モデリングの予測性能を評価した。統合項目を使った際のPR曲線を図7.13に示す。また、代表的な予測性能指標について、統合項目を使っていない結果（7.4.1項の結果）と統合項目を使った結果の比較を表7.8に示す。表7.8の比較から、その予測性能に大きな差異はないことが分かる。これは、今回選んだ統合項目のペアが、重篤事態の予測においては概ね類似した情報を持っているためと考えられる。

XGBoostによる重篤な虐待事例の予測精度（PR曲線）

曲線下面積 = 0.887

閾値0.5とした時の予測精度： Accuracy = 0.616, Precision = 0.854, Recall = 0.614

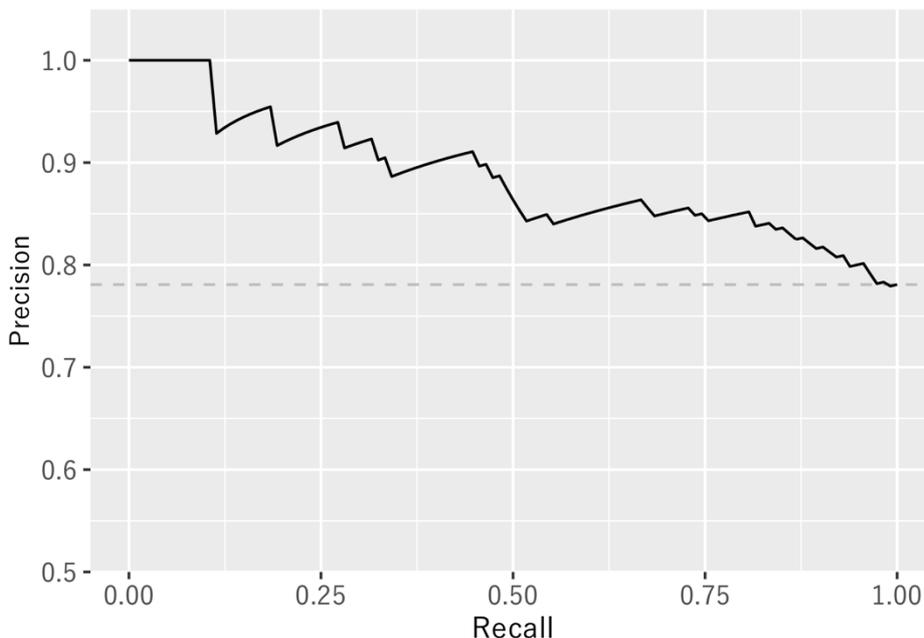


図7.13 XGBoostによる重篤な虐待事例の予測精度（PR曲線）

表7.8 統合項目の有無における予測性能指標の比較

閾値0.5とした時の予測精度（重篤事態の予測）						
統合項目の有無	AUCPR	Accuracy	Precision	Recall	Specificity	Neg.Pred.Value
無	0.876	0.630	0.857	0.632	0.625	0.323
有	0.887	0.616	0.854	0.614	0.625	0.312

続いて、7.4.3項と同様の手順で各セーフティアセスメント項目についてGlobal SHAPとSigned Global SHAPを算出した。表7.9にGlobal SHAPの値が大きい順に上位35項目を列挙した。項目番号3と5の統合項目は順位33、項目番号43と68の統合項目は順位4に位置づけられている。それ以外の項目に関しては、7.4.3項で示した順位とくらべて、多少の順位変化はあるが大きな変化はない。このことから、セーフティアセスメントツールの作成に際しては、これらの項目を統合してもよいと考えられる。

表7.9 統合項目を使った重篤事態の予測におけるGlobal SHAP 上位35項目

順位	項目番号	厚労省区分	統合項目名	Global.SHAP	Signed.Global.SHAP
1	1	社会関係	支援/介入の困難/資源不足	0.219	0.219
2	16	養育者の様子	説明の回避/説明内容の疑念	0.193	0.193
3	4	児童の様子	児童の情緒的問題/対人距離/愛着課題がある	0.140	0.140
4	43and68	統合項目	児童が帰宅を嫌がる・拒否する/児童の帰宅不安/恐怖	0.132	0.132
5	2	養育者の様子	過剰なしつけ/体罰での暴力/正座等の強制/暴力のほめかし	0.121	0.121
6	15	養育環境/生活状況	養育負担の偏り/夜間監護がない	0.086	0.085
7	19	養育環境/生活状況	不適切な養育環境	0.075	-0.075
8	20	家族(きょうだい)情報/世帯情報	経済不安/就労不安定	0.074	0.072
9	11	児童の様子	養育者に対する挑発/エスカレートする行為	0.069	0.021
10	13	養育者の様子	精神科既往歴/不安定/判断力の減退/養育困難	0.059	0.059
11	59	養育者の様子	支援者への攻撃性	0.043	0.043
12	108	その他	今までに経験したことのない事例	0.042	0.042
13	18	児童の様子	摂食/排泄の異常/喘息やアレルギー	0.040	0.040
14	37	養育者の様子	通告による傷つき/プレッシャー	0.040	0.040
15	48	重篤項目	児童の身体に打撲痕や内出血などの外傷がある	0.036	0.036
16	9	養育者の様子	怒りや突発的事態への対処困難	0.033	0.018
17	10	養育者の様子	児童の育てにくさ	0.027	-0.001
18	29	養育者の様子	虐待の黙認/擁護/認識欠如	0.026	0.026
19	98	重篤項目	児童自身が保護・救済を求めている	0.023	0.023
20	24	家族(きょうだい)情報/世帯情報	夫婦間葛藤/対立/話し合い困難/立場が対等ではない	0.023	0.023
21	21	児童の様子	噛み傷/わずかな傷/説明されない傷	0.022	0.022
22	96	初期情報	受賞起点からの時間経過	0.021	0.021
23	8	児童の様子	養育者を過剰に支持する	0.021	0.021
24	28	その他	複数種別の虐待が併発/混合	0.020	-0.020
25	7	児童の様子	過去の心理的/身体的虐待歴	0.019	-0.015
26	6	養育者の様子	育児スキルの不足/不履行	0.018	0.010
27	87	妊娠/出産	若年出産	0.017	0.017
28	34	養育者の様子	人前での暴言暴力/泣いてもあやさない様子	0.017	-0.017
29	14	養育者の様子	養育者の家庭外ストレス	0.017	-0.015
30	26	家族(きょうだい)情報/世帯情報	家族構成/同居人変化	0.016	-0.016
31	12	養育者の様子	支援の拒絶/回避/無関心/支援者への態度に一貫性がない	0.016	0.000
32	75	重篤項目	養育者が、現状解決されていない養育上の課題に対する解決方法を求めている	0.013	0.013
33	3and5	統合項目	養育者の都合が養育より優先/生活が自己中心的/態度から事態改善が見込まれない	0.012	0.012
34	32	児童の様子	学業上での課題を抱える	0.012	-0.012
35	44	養育者の様子	物質/行為依存/発達障害の診断/疑い	0.010	-0.010

7.4.6 採用項目数別の予測性能指標（重篤事態の予測）

7.4.1項では131項目のセーフティアセスメント項目から重篤事態を予測するXGBoostを構築した。さらに、7.4.3項では予測的貢献度を測る指標としてGlobal SHAPを項目ごとに評価した。これらをふまえて本項では、Global SHAPの値が大きい上位数項目だけを用いて重篤事態を予測した際に、その予測性能が影響を受けるかを精査した。

Global SHAPの値が大きい順に採用した項目数（採用項目数）を変えてXGBoostを学習し、その際の精度（Accuracy）を評価した（図7.14）。ここで、XGBoostの学習は7.4.1項で述べた手順で行った。

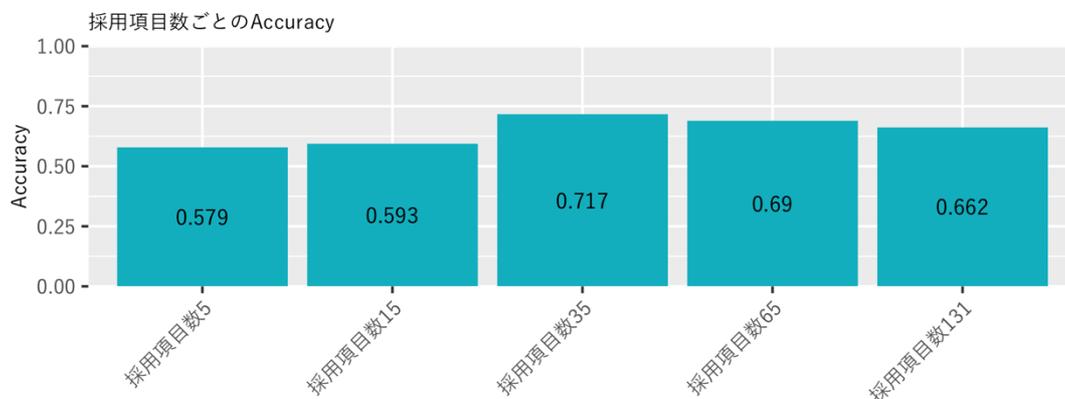


図7.14 採用項目数を変えた際のAccuracy（精度）

図7.14から、採用項目数を35まで減らしても、Accuracyに大きな差はないことが分かる。この結果は、131のセーフティアセスメント項目のうち、Global SHAPの大きな上位数項目が特に予測に貢献していることを示唆する。Accuracy以外の予測性能の評価指標については表7.10に示す。採用項目数を5まで減らすと、RecallとSpecificityのバランスが崩れ、重篤事態の回答データに過学習してしまっていることが見て取れる。

表7.10 採用項目数を変えた際の予測性能指標

採用項目数ごとの予測性能指標						
採用項目数	AUCPR	Accuracy	Precision	Recall	Specificity	Neg Pred Value
5	0.898	0.579	0.325	0.735	0.532	0.868
15	0.903	0.593	0.329	0.706	0.559	0.861
35	0.911	0.717	0.431	0.647	0.739	0.872
65	0.913	0.690	0.404	0.676	0.694	0.875
131	0.910	0.662	0.373	0.647	0.667	0.860

第8章のセーフティアセスメントツールの項目選定に際しては、ここでの解析結果を参照する。

7.4.7 除外項目の検討

有識者より、「項目番号43、68、98は類似の質問内容であり、かつそれらに該当した場合は重篤事態に該当する可能性が非常に高い。このことから、それらの項目はセーフティアセ

メント項目から除外できるのでは。」という指摘があった。そこで、本項では、項目番号43、68、98を除いて機械学習モデルを構築した場合に、その予測精度に影響があるかを精査した。これらの項目を除外した場合とそうでない場合の予測性能指標を表7.11に示す。どちらの場合も、Accuracy、Recall、Specificityが概ね0.6～0.7の値をとっていることから、項目除外による影響は小さいと言える。

表7.11 類似項目を除外した場合の重篤事態の予測性能比較

項目番号 43,68,98 を除いた際の予測性能指標						
除外項目の有無	AUCPR	Accuracy	Precision	Recall	Specificity	Neg Pred Value
無	0.910	0.697	0.4	0.588	0.730	0.853
有	0.911	0.683	0.4	0.706	0.676	0.882

第8章 セーフティアセスメントツールの構成

【第8章の概要】

本事業では、データから得られた各アセスメント項目の信頼性及び妥当性の検証結果をもとに、項目を選定し、アセスメントツールを構成する。本章では、第7章までに得られた解析結果を踏まえ、セーフティアセスメントツールの構成プロセスを述べる。

8.1 目的

第7章までの分析で、アセスメント項目の信頼性（評定者間一致性）・予測的妥当性が検証された。本章では、アセスメントツール構成に向けて、有用な情報が最大限取りこぼされることなく、かつ最小限の項目数となるようアセスメント項目を選定する。

8.1.1 セーフティアセスメントツールの対象と活用場面

本ツールが対象とするリスクは、初動対応段階における死亡リスク、及び致命的行為の並存リスクである。ツールによって判断されるのは、一時保護の要否、あるいは市区町村から児童相談所への通告/送致の要否であり、重篤事態の見落とし防止・検知力の向上を目指す。したがって、その活用場面は「通告から一時保護(または児相への通告/送致)の要否判断を経た援助方針会議まで」、および「在宅支援継続中に現状方針を見直し、一時保護の要否を再検討する場面」となる。

本セーフティアセスメントツールが想定する活用場面・ターゲット

誰が：児童相談所職員と市区町村職員

いつ：通告～援助方針会議まで

またはリミットアセスメント*を経て、一時保護の要否を再検討する場面

何を：重篤な項目に該当があるかどうか確認する

何のために：一時保護(児相)、または通告/送致(市区町村)の要否判断

(※リミットアセスメントとは：現在の支援方針を見直すための基準・条件を設定し、それを満たしているかを判断するためのアセスメント。在宅支援継続中に子どもの安全が確認できなくなるなど、現状の支援方針を見直す必要があると判断された場合、再度セーフティアセスメントを行い、児童相談所による一時保護、または市区町村による児童相談所への通告/送致の要否判断を行う。)

8.1.2 セーフティアセスメントツールの形式

全てのアセスメント項目に該当・非該当を入力するのが理想的だが、調査を経ていない段階では、該当するかどうか判断できない項目も多い。そのため、調査を進めていきながら、新しい情報が判明次第、つどアセスメントツールへ入力していく形が望ましい。特に、web