

4. 補 データクリーニングについて

今回の調査は、調査概要で説明したように、調査会社の保有する調査モニターを用いて回答者をリクルートし、パソコン回答を基本としてインターネットで回答してもらう方法をとっている（これらの方法を総称して、ここでは「インターネット調査」と表す）。調査会社のモニターを利用するサンプリングの方法や、インターネットでの回答を得ることの社会調査上の意義や方法的な問題点についてはすでに指摘されており（たとえば三浦・小林 2015 など）、本調査においても、データ確認作業の中で、回答項目間の論理矛盾や、マトリックス形式における盲従化回答など、疑義をもつケースの検討に時間を要した。ここでは、作業チームが行ってきたデータクリーニング作業の記録をまとめるとともに、女性の妊娠経験や出生前検査というテーマの社会調査における課題を整理しておく。

今回の調査では、妊娠経験のある女性のみが対象であるが、出生前検査という一般的とは言えない検査の理解や経験について尋ねた。一連の回答をみていると、（妊娠経験があっても）出生前検査にあまり知識がなかったり、自身が経験した検査と、調査で尋ねている検査を勘違いしている（たとえば、単純な経過観察としての採血を NIPT と誤解している）と思われるような回答者もいる。また、一番直近の妊娠から時間的経過も、現在妊娠中から十数年前まで回答者によって様々であり、回答者の記憶の状態や出生前検査の条件等も異なる。そのため、特に出生前検査の受検に対する回答については慎重にクリーニングする必要があると考えた。

クリーニング作業はインターネット調査を担当する作業班（田中・菅野・井原）が担当し、データの大幅な修正は田中が行った。最初に、一番最近の妊娠時期が不明だとその後の分析が難しくなるため、一番最近の妊娠時期について「答えたくない」という回答、21名は除外し、2357名を有効回答とした。その後、作業班のメンバーから指摘された論理矛盾のある回答箇所は、研究会メンバー全員で確認した上で、必要に応じて回答を修正し、場合によっては分析対象からの除外、もしくは無回答や非該当への変更を行った。修正箇所は表1参照。

また、出生前検査に関する質問項目（予備調査の検査受検など）において、この後に述べる3つの条件のいずれかにあてはまる回答136名（全体の5.7%）は、決定的な不良回答として有効回収から削除するまでには至らないものの、受検の有無の回答自体に疑問が生じるものとみなし、出生前検査関連の集計の対象外とした。なお、この136名は検査に関する質問項目以外の回答では論理矛盾等が認められなかったことから、検査に関する質問項目に限って集計の対象外とし、データとしては全体数からは削除しなかった。同様に、妊娠や出産に関する意識など、主にマトリックス形式で回答を得ている質問群では、いずれの質問にも「答えたくない」が出現しない者に限定して集計した。そのため、本報告では質問ごとのサンプルサイズが異なっている。以下では、出生前検査のデータクリーニングについて詳細を述べる。

まず、NIPT は世界においても 2012 年から提供された検査だが、2011 年以前に NIPT を受けたとする回答 18 件、および NIPT を受けたにも関わらず、【新型出生前検査 (NIPT) には流産のリスクがある】との正誤問題に不正解だった回答 30 件については、NIPT を受けたとする回答に疑問が生じたため集計対象外とした。

次に、超音波検査 (43 件)、羊水検査 (1 件)、母体血清マーカー検査 (1 件)、NT 検査 (1 件) および着床前診断 (1 件) について、調査項目間で受検の有無の回答に明らかな矛盾が認められた。例えば、自由記述欄に検査を“受けていない”と記載されているのに、受検の有無を尋ねる選択形式の質問では“受けた”と回答があるというようなケースである。これらの回答についても同様に集計対象外とした。

最後に、6 種類 (超音波検査、羊水検査、母体血清マーカー検査、NIPT、NT 検査および着床前診断) の検査の受検の有無を尋ねた 6 つの質問全てにおいて、2357 件中 2079 件 (88.2%) の回答は、「検査を受けた」、「検査を受けなかった」、「他の医学的検査から受ける必要はなかった」、「妊娠時に検査が実施されていなかった」のいずれかだったのに対し、「わからない」、「覚えていない」、「答えたくない」、もしくは非該当処理した回答 (超音波検査の受検の有無で予備調査と本調査で回答に齟齬があったもの 表 1 参照) が 1 つ以上ある回答は 278 件、2 つ以上は 95 件、3 つ以上は 57 件だった。278 件については検査経験の記憶が不確かな可能性や、検査名を知らなかったり誤解している可能性が懸念された一方で、検査名を知らなかったり覚えていない可能性もありえることから、わからない等とする回答が 2 つ以下の場合を残し、半数にあたる 3 つ以上該当する場合 (57 件) のみ分析対象外とすることとした。したがって、出生前検査に関する項目の集計結果は、全回答者のうち検査の記憶が相対的にあいまいな集団が除外されたサブサンプルの数値であることに留意する必要がある。

なお、6 種類全ての検査を「受けた」という回答は全体で 28 件あり、上記の条件で検査の集計対象から除外しても 9 件残ったことから、この 9 件の回答の信憑性をはかる基準を議論した。その結果、検査の受検基準が厳格である着床前診断に関する情報を参照することにした。日本産科婦人科学会の倫理審査を受けて実施されている着床前診断をこれだけ多くの人 (本調査全体で「受けた」という回答があったのは 54 名、「受けた」/ [「受けた」+「受けていない」] で算出した受検率は 2.7%) が、「受けた」と回答しているとは考え難いため、着床前診断受検者 54 名について、日本産科婦人科学会の当該倫理審査の基準にあたる、自然流産・自然死産の経験回数、自分もしくは近親者 (親、子、親族) に障害のある人がいるかどうか、および配偶者の有無の観点から、調査内の回答をさらに精査した。しかし、着床前診断を受けたことが疑われるような十分な条件が見いだされなかったため、この 9 件については集計対象に残した。

表 1 出生前検査に関する内容とクリーニングの対応

項目	クリーニング内容 ※右の件数は、修正箇所数。該当する ID の件数ではない。	件数
超音波検査受検	<p>・「その他の検査」の受検の「具体的に」の欄に”エコー”など振り替え可能な検査名の記載あり →その他の検査を「受けた」を「受けていない」へ振替、具体的な記入を非該当へ。</p> <p>・「その他の検査」に「胎児ドックとじゅうもう」と記載あり →胎児ドックは超音波検査に振り返られるため、削除。じゅうもう検査のみ残す。</p>	93
	<p>・「その他の検査」の受検の「具体的に」の欄に「エコー」と記載あり →【①超音波検査の受検回数、「超音波検査を受けていない」(1件)、②超音波検査を受けたか「超音波検査を受けたかわからない」(1件)、③超音波検査を受けたか「超音波検査を受けていない」(4件)】を「受けた」へ。「その他の検査の受検」を「受けた」を「受けていない」へ振替。</p>	6
	<p>「不妊検査や治療の経験」で「はい」(不妊治療を受けた)ことから超音波検査も受けたと考える→超音波検査を受けたか「超音波検査を受けていない」を「新コード」へ(9件)、「超音波検査を受けたかわからない」(5件)と②「超音波検査を受けていない」(9件)を「受けた」へ振替。</p>	23
	<p>「妊娠中に胎児の性別がわかったか」で「超音波画像で胎児の性別がわかった」ことから超音波検査も受けたと考える→超音波検査を受けたか「超音波検査を受けていない」を「受けた」へ振替。</p>	34
	<p>「羊水検査を受けたか」で「羊水検査を受けた」ことから超音波検査も受けたと考える→超音波検査を受けたかで「超音波検査を受けたか覚えていない」を「受けた」へ振替。</p>	17
超音波検査わかったこと	<p>「超音波検査でわかったこと」で「超音波検査でわかったこと「その他」の具体的記述欄に「受けていない」と記載あり→当該質問の回答を無回答に振替。</p>	6
超音波検査希望	<p>「誰が超音波検査を希望したか」の複数回答で、「自分で超音波検査を希望した」と「超音波検査を希望しなかった(断った)」を共に選択し回答に齟齬がある。ともに「誰が超音波検査を希望したか」と「超音波検査を受けた回数」は受けたと回答。→「誰が超音波検査を希望したか」の「希望した」の回答は残し、「断った」は無回答へ振替。</p>	4
超音波検査理由	<p>「超音波検査を受けたか」と「超音波検査の回数」で共に「受けた」とあるのに、「超音波検査を受けなかった理由」に回答あり。【※「超音波検査を受けた理由」は無回答。「超音波検査でわかったこと」にも回答あり】</p>	41

	→「超音波検査を受けなかった理由」の複数回答の回答を無回答へ振替。	
	「超音波検査を受けたか」と「超音波検査の回数」で共に「受けていない」とあるのに、「超音波検査でわかったこと」に回答あり。【※「超音波検査を受けなかった理由」に回答あり】→「超音波検査でわかったこと」を無回答へ振替。	28
	「超音波検査を受けたか」で「わからない」「覚えていない」、また、「超音波検査を受けた回数」で「受けていない」の場合、受けていないと考える。→「超音波検査を受けた理由」と「超音波検査でわかったこと」はいずれも無回答へ。【※「超音波検査を受けた理由」と「超音波検査を受けなかった理由」への回答はなかったため修正不要】	6
超音波検査受検：非該当	「超音波検査を受けたか」で「受けていない」か「答えたくない」、「超音波検査を受けた回数」で受けた回数を回答している場合、齟齬があるので非該当→両質問を非該当へ。「超音波検査を受けた／受けない理由」と「超音波検査でわかったこと」も非該当へ振替。	190
	「超音波検査を受けたか」で「受けた」、「超音波検査を受けた回数」で「受けていない」の場合、齟齬があるので非該当→「超音波検査を受けたか」と「超音波検査を受けた回数」を非該当へ。「超音波検査を受けた理由」と「超音波検査でわかったこと」も非該当へ振替。	156
	「超音波検査を受けた回数」で「答えたくない」なのに、「超音波検査を受けたか」で「答えたくない」以外の回答で、齟齬がある→「超音波検査を受けたか」と「超音波検査を受けた回数」を非該当へ。「超音波検査を受けた／受けない理由」と「超音波検査でわかったこと」も非該当へ振替。	155
NT 受検	自由記述の「妊娠経験のお気持ち」に「NT 検査」と記載されていた。→「NT 検査を受けたか」を「受けていない」から「受けた」へ振替。	1
NIPT 希望	「NIPT を受けたか」で「受けていない」だが、「NIPT の受検経緯」は、NIPT を「希望しなかった」と「希望した」の両方が選択されていた。→「受けていない」と判断することとし、NIPT を「希望しなかった」を残し、「希望した」を非該当へ振替。	1
マーカー受検希望	「母体血清マーカー検査を受けたか」で「受けた」だが、「マーカーの受検経緯」はマーカーを「希望した」と「希望しなかった」の両方が選択されていた。→マーカーを「希望した」を残し、「希望しなかった」を非該当へ振替。	2

	「母体血清マーカー検査を受けたか」で「受けていない」だが、「マーカーの受検経緯」はマーカーを「希望した」と「希望しなかった」の両方が選択されていた。→マーカーを「希望しなかった」を残し、「希望した」を非該当へ振替。	2
マーカー受検	「マーカー検査の結果」の「その他：具体的に」と「マーカー検査結果後の行動」の「その他：具体的に」で「受けていない」と記載があった。→前者の記載を非該当、後者の記載を「回答なし」へ修正。加えて、「マーカー検査受検有無」を「受けた」から「受けていない」へ修正。「マーカー検査結果後の行動」は、「その他」を非該当とし、「マーカー受検理由」は、「上記のいずれもあてはまらない」を非該当へ振替。	6
羊水検査受検	「羊水検査を受けたか」で「受けた」と回答あるが、「羊水検査を受けた後の対応」の「その他：自由記述」で、「受けていない」と回答があった。→「羊水検査を受けたか」を「受けていない」へ振替。「羊水検査を受けた後の対応」で「その他」は非該当へ。「羊水検査を受けた後の対応」のその他「受けていない」と自由記述を記述なしへ修正。「羊水検査でわかったこと」の選択肢「何もわからなかった」は、受けなかった人が選択できる回答なので、そのまま残す。「羊水検査を受けた理由」の選択肢「上記のいずれもあてはまらない」を非該当へ。「羊水検査を希望したか」を「自分で希望した」を非該当へ振替。	5
羊水検査希望	「羊水検査を受けたか」で「受けた」だが、「羊水検査の受検経緯」は羊水検査を「希望した」と「希望しなかった」の両方が選択されていた。→羊水検査を「希望した」を残し、「希望しなかった」を非該当へ振替	1
着床前診断受検	「その他の検査の受検」で「超音波検査、着床前診断」と記載されているのに、「着床前診断を受けたか」は「受けていない」と齟齬がある→「着床前診断を受けたか」を「受けた」へ振替。	1
年齢	「現在妊娠されていますか」で「はい」で、現在の年齢が1番最近の妊娠年齢より2歳以上低い→1番最近の妊娠年齢を非該当へ振替。	4
	「現在妊娠されていますか」で「いいえ」で、現在の年齢が1番最近の妊娠年齢より1歳以上低い→1番最近の妊娠年齢を非該当へ振替。	3
胎児の性別	「性別を知った方法」で「その他：具体的に」で、中絶や流産で妊娠が継続しなかったと記載あり→選択肢「妊娠が継続しなかったためわからなかった」に振替。	18
	「性別を知った方法」について選択肢「その他」で、超音波検査以上の検査で分かった、という内容は新変数として設定。	8

	「性別を知った方法」について選択肢「その他」で、「医師から説明された」という記載がある→選択肢「医療者から説明されてわかった」に振替。	4
	「性別を知った方法」について選択肢「その他」で、医師に聞かなかったと記載がある→選択肢「妊娠中にはわからなかった（聞かなかった）」に振替。	12
妊娠感覚	「妊娠したと感じた時」で「その他」の自由記述を選択肢内の回答に振替。	10
検査薬	「妊娠したと感じた時」で選択肢「市販の妊娠検査薬で」と「妊娠検査薬の利用」の齟齬→非該当に。	14
35歳	「一番最近の妊娠の時期」と「妊娠がわかった時の状況」の選択肢「35歳以上であった」が一致しない→「妊娠がわかった時の状況」を無回答に修正。	14
不妊治療	「不妊治療の経験」では「なし」なのに、「一番最近の妊娠での不妊治療」では「あり」となっている→妊娠中のため、経験に入っていないことを考慮し、「不妊治療の経験」を「あり」に変更。	3
流産	流産経験があり、「超音波検査を受けたか」で「受けなかった」と回答→「受けた」に振替。	1
	「超音波検査を受けた理由」で「前の妊娠が流産・死産だったから」に回答があり、流産・死産をしたと思われる→何回したかはわからないため、「自然流産・自然死産の経験回数」を「x」とした。	6
	羊水検査の「その他」の「具体的に」の記述内容から、選択肢「障害の有無」に振替。	2
病院	「出産した場所」の「その他」の記述内容から、選択肢「総合病院・総合病院」に振替。	2
子	子どもの出生順位が逆転→修正	2
	子どもの年齢・性別→論理整合性に欠ける回答を非該当に。	5
子どもの性別	「子どもの性別」で選択肢「妊娠中でまだわからない」が複数回回答されている→非該当に。	52
	「子どもの性別」で選択肢「妊娠中でまだわからない」が複数回、回答されている→ずれた回答をしたと解釈し、「妊娠中でまだわからない」を「女」に振替。	1

自分と夫の仕事	自身の仕事の「その他」の具体的記述を、内容的に妥当な選択肢がある場合振替。	52
	配偶者の仕事の「その他」の自由記述から、配偶者の学歴に関する質問の回答を「配偶者はいない」に振替。	12
	配偶者の仕事の「その他」の自由記述で「パートナーはいない」とある→非該当に振替。	12
	配偶者の仕事の「その他」の具体的記述を、内容的に妥当な選択肢がある場合、振替。	85
	本人の年収と世帯年収を比較して世帯年収のほうが多いように修正。例えば、本人が 600~800 万で、世帯がなしと回答されている場合、本人をなしへ、世帯を 600~800 へ修正した。	8

以上のように、出生前検査に関する質問項目群については、同一サンプル内での回答の論理矛盾や、回答者の理解ちがいなどを検討し、必要に応じてデータの修正を加えている。

インターネット調査では、例えば直近の妊娠年齢を尋ねる場合も、妊娠開始年と妊娠終了年の別に尋ね、プルダウン方式の選択肢の提示や、一覧形式で数字を入力してもらう方法など、質問紙調査等の他の調査とくらべ、容易に正確な情報を捕捉できるというメリットがある。しかし、調査モニターの中にはもっぱら回答によるポイント（謝礼品）獲得のために、どのような調査でも応諾し、最小限の時間やコスト（全部「3」を選ぶなど）で回答したり、悩む場合にはあえて選択せずに「わからない」や「答えたくない」といった選択肢を選ぶなどの態度が取られている可能性がある（なお、回答時間が早すぎる場合や、すべて「わからない」と回答している場合には、一般的にはデータから除外される）。そのような「ノイズ」をどのように定義するのか、またデータや分析から排除するか否か、とりわけ「わからない」という選択肢が重要な意味をもつ本調査のようなテーマにおいては慎重に検討する必要があるだろう。また本調査で採用した対象者の抽出条件によるバイアスをどのように考慮するのも重要な検討課題である。本調査の記録が、今後の様々な調査や研究の参考になればと願っている。

参考文献

三浦麻子・小林哲郎, 2015, 「オンライン調査モニタの Satisfice に関する実験的研究」『社会心理学研究』 31-1, 1-12. http://doi.org/10.14966/jssp.31.1_1