

令和 2～3 年度 食品健康影響評価技術研究 研究成果報告書（終了時）

研究課題名 (研究項目名)	食品用器具・容器包装に用いられるビスフェノール A のリスク評価に資する科学的知見の検討に関する研究（課題番号：JPCAFSC20202007） (4 その他 (2) その他食品健康影響評価に資する研究・調査)
主任研究者	研究者名：青山 博昭 所属機関：一般財団法人残留農薬研究所

I 研究期間及び研究目的等

1 研究期間

令和 2 年度～令和 3 年度（2 年間）

2 研究目的

食品安全委員会では、平成 22 年 7 月に開催された第 13 回器具・容器包装専門調査会で、生殖発生毒性等に関するワーキンググループから「ビスフェノール A (BPA) に関する健康影響について：中間取りまとめ」が報告された。しかし、中間取りまとめの時点では低用量影響に関する評価が定まっておらず、耐容一日摂取量 (TDI) の設定が極めて困難であるとの理由から、低用量影響に関する新たな科学的知見が得られた時点で最終的な評価を取りまとめることとされた。

その後、BPA の低用量影響に関する科学的知見が次第に集積されつつあることを考慮した食品安全委員会は、BPA に係る評価手法及び科学的知見に関する調査を令和元年度に行い、結果を取りまとめた。この調査では、BPA の体内動態、毒性、ばく露量及び疫学調査の結果を報告した公表論文を網羅的に検索し、約 280 文献（ヒト研究に関する論文を約 50 文献、動物実験に関する論文を約 230 文献）を収集し、その概要を整理した。しかし、未だ有用性を検討すべき文献が数多く残されており、それらの収集・整理が喫緊の課題とされた。

本研究は、上述の如き状況を鑑みて、先の調査方針を継承しつつ、BPA のリスク評価に向けてさらにその内容を充実させることを目的とした。すなわち、有用性を評価すべき最大約 800 文献を対象として可能な限りこれらの文献を収集し、内容に応じて体内動態調査、毒性評価、ばく露量評価、及び疫学調査の 4 分野に整理することを第一の目的とした。また、ヒト研究及び動物試験についてはそれぞれ信頼度を評価するための統一的な基準を設定して、BPA のリスク評価に用いるべき信頼性の高い文献を選抜することを第二の目的とした。さらに、動物実験分野の文献について、リスク評価に資するようエンドポイント別に信頼できると判断された文献の内容を精査して、リスク評価のどの場面でどのような内容を用いるべきかを提言することを第三の目的とした。

3 研究体制

研究項目名	個別課題名	研究担当者（所属機関）
最新のBPA関連文献の収集	文献収集とデータベースの構築	青山博昭 （残留農薬研究所） 井口泰泉 （横浜市立大学）
動物実験分野におけるBPA関連文献の信頼性評価と、信頼性が高いと判断されたデータの抽出	文献一次選抜基準の設定と一次選抜の実施	青山博昭 （残留農薬研究所）
	文献信頼性評価基準の設定	井口泰泉 （横浜市立大学）
	文献信頼性評価及び主要データの抽出	石井雄二 （国立医薬品食品衛生研究所） 北條仁 （残留農薬研究所）
	エンドポイント別のデータ整理	井口泰泉 （横浜市立大学）
疫学分野におけるBPA関連文献の信頼性評価	文献等情報の妥当性の検討	祖父江友孝 （大阪大学）
	ヒトへの健康影響に関する証拠の信頼度の検討	八谷寛 （名古屋大学）
最新の動物実験結果に基づく提言の取りまとめ	BPAリスク評価への提言	青山博昭 （残留農薬研究所） 井口泰泉 （横浜市立大学）

4 倫理面への配慮について

本研究では、文献調査、文献の信頼性評価基準の設定、信頼できる文献の選定、及びBPAのリスク評価に資する文献情報の抽出が主な業務であり、研究対象に対する倫理的な問題は発生しないと想定される。また、一般財団法人残留農薬研究所には2名の内部委員（常務理事及び業務執行理事）と1名の外部委員からなる利益相反委員会が、横浜市立大学には学内規定に沿った利益相反委員会がそれぞれ設置されており、主任研究者及び分担研究者（所属機関及び家族を含む）のいずれにも、BPA またはBPAが残留する可能性のある樹脂製品を製造・販売する会社あるいは関連団体から資金の提供を受けておらず、開示すべき利益相反はない。

II 研究内容及び成果等

1 研究項目：最新のBPA関連文献の収集

(1) 個別課題：文献収集とデータベースの構築

(青山博昭(残留農薬研究所)、井口泰泉(横浜市立大学))

食品安全委員会が令和元年度に実施した調査事業では、約280文献を収集して、概要が取りまとめられた。本研究では、引き続き文献検索を進めた結果、約770文献を追加収集することができた。そこで、これらの作業で収集された合計1,052文献をデータベース化し、動物実験分野及び疫学分野の文献については、それぞれについて文献ごとに概要を記載して、信頼性評価の基礎資料として整備した。

2 研究項目：動物実験分野におけるBPA関連文献の信頼性評価と、信頼性が高いと判断されたデータの抽出

(1) 個別課題：文献一次選抜基準の設定と一次選抜の実施

(青山博昭(残留農薬研究所)、井口泰泉(横浜市立大学)、石井雄二(国立医薬品食品衛生研究所)、北條仁(残留農薬研究所))

信頼性評価に先立って、まずヒトへの外挿性に係る基準に沿って文献の妥当性(Relevancy)を評価して有用な文献を抽出することとし、以下の基準を設定した。

- 1) 文献から得られる情報に基づいてリスク評価を実施するためには、ビスフェノールA(BPA)を経口投与(強制経口投与、混餌投与または飲水投与)した実験であること
- 2) 一般的な動物種及び系統の無処置動物を用いた実験であること(遺伝子改変動物や去勢動物を用いた実験ではないこと)
- 3) 対照群と3用量以上のBPA投与群が設定された実験であること
- 4) BPAの低用量影響や非線形用量反応関係の有無が検討できるよう、一般的なリスク評価においてTDI(Tolerable Daily Intake)設定時にPOD(Point of Departure)として採用されている5 mg/kg bw/day以下の用量(混餌投与または飲水投与実験にあつてはその近傍の用量)を含む用量設定がなされた実験であること

次いで、協力研究者(石井雄二及び北條仁)の協力を得て、BPAの代謝あるいは体内動態を解析した結果を報告した文献を除く動物実験分野の文献約400報を上記(1)に示した基準に沿って評価し、信頼性を詳細に評価すべきと判断された文献147報を選抜した。

(2) 個別課題：文献信頼性評価基準の設定

(青山博昭(残留農薬研究所)、井口泰泉(横浜市立大学))

上記(1)で選抜された文献の信頼性を評価するため、研究体制、材料及び方法、並びに結果に関して適切な評価項目と評価基準を設定することとし、それらを別添資料表1に示す如く整理した。その際には、EFSAプロトコール(Bisphenol A (BPA) hazard assessment protocol (EFSA Supporting publication, 2017:EN-1354))やOHATハンドブック(Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration (Office of Health Assessment and Translation (OHAT), 2019))で提唱された信頼性評価の考え方(Risk of Bias Tool)も参考にした。

(3) 個別課題：文献信頼性評価及び主要データの抽出

(石井雄二(国立医薬品食品衛生研究所)、北條仁(残留農薬研究所))

上記(1)で選抜された147文献について、(2)で設定された基準に沿ってそれらの信頼性を詳細に評価した。信頼性評価作業は、二段階(一段階目「研究体制」並びに「材料及び方法」、二段階目「結果」)に分けて実施した。まず、「研究体制」及び「材料及び方法」の2項目についての評価を実施して、結果を集計した。集計に当たっては、いずれか一方の評価者の評価結果において、(a) 評定平均値が「3.0」以上であって個別の評定値として「1」と評価した項目がなかったこと、及び(b) 危害分析を目的とした実験であるとみなし得ると評価したこと、の2条件を満たしていた場合に、信頼に足る文献と判断した。その結果、36文献が抽出されたことから、これらについて引き続き「結果」の信頼性を評価することとした。これらの作業により、「結果」の項についても2名の評価者の評定平均値が「3.0」以上であった14文献を、信頼性が高いと評価した。また、いずれか一方の評価者が信頼性を認めた(評定平均値が「3.0」以上であった)7文献については、担当者間で再度協議して、さらに2文献を信頼性が高いと判断した。

上記の信頼性評価により十分に信頼できると判断された合計16文献(別添資料表2)を対象としてそれらの内容を精査し、対照群とBPA投与群の間で統計学的に有意な差が検出されたエンドポイントをすべて抽出して整理した。その結果、3文献(データベース上の文献番号:119、171及び794)では、いずれの検査項目にもBPA投与の影響が検出されていなかった。また、残る13文献では何らかの指標に対照群とBPA投与群の間で差が観察されていたものの、8文献(文献番号:11、17、80、164、608、696、698及び729)については、報告されたそれらの結果(一般的な毒性指標とは言い難い、肝脂肪、血清アポプロテイン、nNOS発現等の僅かな変化、エストロゲンに対する感受性の系統差の検出、高脂肪食あるいは発がん物質との併用により群間に生じた僅かな差の検出等)を毒性評価に用いることに疑問が残り、現時点では検出された対照群とBPA投与群との差を悪影響と解釈することは困難であると結論された(別添資料表3)。したがって、今後のBPAのリスク評価に有用な信頼性が高いデータを提供し得ると判断された文献は、5報(文献番号:132、133、192、750及び776)に限定された。これらの文献は、5 mg/kg bw/day未満の用量を含む様々な用量のBPAを母動物の妊娠期間(文献番号:192、750及び776)または2世代あるいは3世代にわたって交配前8または10週間と妊娠・哺育期間(文献番号:132及び133)に投与して、母動物と児動物への影響を最長2年間にわたって詳細に調べた結果を報告したものであった。これらの文献から抽出されたデータを、別添資料表4に示す。これらのうち、3文献(文献番号:132、133及び192)では、50 mg/kg bw/day以上の高用量群でのみBPA投与の影響が観察されており、非直線的な用量反応関係は認められていなかった。一方、残る2文献(文献番号:750及び776)はいずれもが米国において実施されたBPAの低用量影響検証実験(CLARITY-BPA Study)の結果を報告したものであり、0.0025~25 mg/kg bw/dayの範囲の用量で様々な指標に対照群との間で統計学的有意差が検出されていた。また、観察された結果をみると、それらの多くは用量反応関係を欠くものであった。

(4) 個別課題：エンドポイント別のデータ整理

(井口泰泉(横浜市立大学))

上記(3)で、今後のリスク評価に有用で信頼性の高い情報を提供し得ると判断された5文献を対象に、対照群とBPA投与群の間で統計学的に有意な差が検出されたすべてのエンドポイントについて、BPAの投与用量と検出された差との関係を示した総括表(別添資料表5)を作成した。

3 研究項目：疫学分野における BPA 関連文献の信頼性評価

(1) 個別課題：文献等情報の妥当性の検討

(祖父江友孝 (大阪大学)、八谷寛 (名古屋大学))

信頼性評価に先立って、収集された文献数を詳細な評価が可能な数にまで絞り込むため、すべての疫学文献を対象として、エンドポイントの信頼度を評価した。すなわち、報告されたエンドポイントが以下に示す4つの分類のいずれに該当するかを判定し、1)、2) または3) に分類された文献を信頼性評価の対象として選抜した。

- 1) エンドポイントが何らかの疾患 (糖尿病, 子宮内膜症, 喘息, がん等) として明確に定義されており, 信頼性を評価するに足ると判断される文献
- 2) エンドポイントが何らかの症状 (停留精巣, 流産, 死産, 早発性思春期等) として明確に定義されており, 信頼性を評価するに足ると判断される文献
- 3) エンドポイントが検査値 (体重, 血糖値, 血圧等) ではあるものの, それらの異常値の定義が明確であることを根拠に, 信頼性を評価するに足ると判断される文献
- 4) パラメータの設定が不適切であったり, 検査値の定義が不明確であったりして, 信頼性評価に値しないと判断される文献

まず、上述の基準に沿って2名の評価者が収集された疫学文献423報の内容を検討し、両者が4)に分類した文献、総説 (メタアナリシスを含む) や症例報告、体内動態に関する文献の計105文献を除外して、ヒトへの健康影響に関する証拠の信頼度を詳細に検討すべき文献318報を選抜した。

(2) 個別課題：信頼性基準の作成と信頼性評価の実施

(祖父江友孝 (大阪大学)、八谷寛 (名古屋大学))

次いで、EFSAのRisk of Bias Tool及びOHATハンドブックを参考に、コホート研究と症例対照研究を対象として、信頼性評価基準を作成した (別添資料表6)。この信頼性評価基準には、先に実施した各論文をグループ分けするエンドポイントの種別も含めた。また、横断研究は信頼性評価の対象とはしなかったため、先に絞り込まれた318文献のうち、111文献を信頼性評価の対象から除外した。さらに、2名の評価者が検討を行い、信頼性評価基準 (別添資料表6) の5 (ばく露測定方法) と8 (エンドポイント測定方法) の得点が3未満であった56文献を除外したところ、合計151報の文献が選抜された。

(3) 個別課題：ヒトへの健康影響に関する証拠の信頼度の検討

(祖父江友孝 (大阪大学)、八谷寛 (名古屋大学))

BPAによるヒトへの健康影響に関する証拠 (エビデンス) の信頼度のランク付けをエンドポイント別 (エンドポイントのクラスター別) に行うため、はじめにこれまでに選択された151の各論文を、その報告しているエンドポイントに基づきグループ化した。具体的には、EFSAが採用した分類に原則としてしたがって、各文献をエンドポイントのクラスターにグループ化した (別添資料表7及び8)。

ヒト健康影響に関する証拠の信頼度を検討するに当たっては、動物実験分野の文献の評価結果から精巣重量及び卵巣への影響に関してヒトにおける評価の必要性が示唆されたこと、及びそれらのエンドポイントのクラスターには信頼度のランク付けを行うに十分な文献数が含まれていたとの事実に基づいて、男性妊孕性、女性妊孕性の2つのエンドポイントのクラスターを選択して、BPAの影響についての証拠の信頼度のランク付けを行った (別添資料表9及び10)。その際には、男性妊孕性のエンドポイントクラスターをさらに男性不妊及び精液の質の2群に、女性妊孕性のエンドポイントクラスターを

さらにIVF（体外受精）結果あるいは妊娠までの期間及び卵巣機能の2群に分け、それぞれ証拠の信頼度のランク付けを実施した。

動物実験分野の文献調査からヒトにおける検討の必要性が考慮されたエンドポイントのクラスターは他にもあったが、これらのクラスターに含まれる疫学文献の数が不十分であったことから、今回は証拠の信頼度のランク付けの検討は実施しなかった。

上述の作業は、以下のAからDのステップから構成された。

- A) 評価対象とするエンドポイントクラスター（その中の群）を構成する文献について、コホート研究が60%以上の場合、そのエンドポイントクラスター（あるいは群）のInitial rating（初期評価）を中等度、それ以外は低いとした。
- B) 証拠の信頼度を低める（Downgrading）要因として、Risk of Bias（文献の信頼性評価が低い文献の割合）、Unexplained inconsistency（研究間の理由のつかない結果の不一致のある文献の割合）、Indirectness（代替指標、代替的对象集団での検討による間接的な知見を報告する文献、外的妥当性がない文献の割合）、Imprecision（推定された関連性指標の精度が低い文献の割合）を採用した。
- C) 証拠の信頼度を高める（Upgrading）要因として、Magnitude（強い関連性を報告する文献の割合）、Dose response（量反応関係があることを報告する文献の割合）、Residual confounding（残余交絡のある文献の割合）、Consistency（異なる研究デザイン・研究対象集団での結果の一致）を採用した。
- D) Downgrading、Upgradingの結果に基づき、初期評価を調整し、当該エンドポイントクラスター（あるいは群）の証拠の信頼度の最終評価とした。その際、具体的な証拠の内容（関連あり、関連なし）を付記した。

ステップB及びステップCでは、証拠の信頼度を低下または上昇させる要因とそれらの評価基準を、以下のように設定した。

証拠の信頼度のDowngradingとUpgradingの基準

Down-grading	<p>(1) 文献の信頼性評価が低い（Risk of Bias：信頼性は、「ばく露測定」と「結果測定」に係る信頼性に基づいて評価する。「ばく露測定」または「結果測定」のいずれかが「2」点以下の時、信頼性が低いと考える）。それ以外の信頼性評価項目のうち、「集団の比較可能性」はDowngradingのindirectness、「対象者数」はDowngradingのImprecision、「交絡の制御」と「統計学的手法」はUpgradingのresidual confoundingと重複するため、文献の信頼性評価の低さを判定する要因とはしなかった。また、「エンドポイント脱落による除外」、「ばく露欠測による除外」、「ばく露指標の妥当性及び対象集団のばく露の変動範囲」、「結果報告」が低い文献はほとんどなかった。</p> <ul style="list-style-type: none"> ∞ 60%以上の文献の信頼性が低い場合 ○ 信頼性が低い文献が40%以上60%未満の時 ±0 信頼性が低い文献は40%未満
--------------	---

Down-grading	<p>(2) 研究間の結果の不一致 (Unexplained inconsistency : 不一致の理由が既知の場合を除く)。層化分析が実施され、特定の層のみで結果が認められる場合、本項目の不一致文献に該当すると考えた。</p> <ul style="list-style-type: none"> ⊖⊖ 60%以上の文献で不一致が認められる場合 ⊖ 結果が不一致の文献数が40%以上60%未満の時 ±0 不一致の文献はあるが、その数が全体の40%未満の時
Down-grading	<p>(3) 代替指標、代替的対象集団での検討による間接的な知見、外的妥当性のなさ (Indirectness : 健康障害としての臨床的意義が不明確な結果変数を用いている、対象集団が結果を適用する集団と本質的に異なる、対照群の設定が妥当でない)</p> <ul style="list-style-type: none"> ⊖⊖ 60%以上の文献で間接的な知見の提示 ⊖ 間接的な知見提示の文献数が40%以上60%未満の時 ±0 間接的な知見を提示した文献はあるが、その数が全体の40%未満の時
Down-grading	<p>(4) 推定された関連性指標の信頼性が低い (Imprecision: サンプルサイズまたはエンドポイントの発生数が小さく、関連性指標であるオッズ比やリスク比の信頼区間が広いこと。オッズ比またはリスク比の信頼区間の下限と上限の比が10を超える場合に広いと判定する。またオッズ比やリスク比が示されていない研究においては、推定値の信頼区間が100以上である、サンプルサイズがコホート研究または横断研究で5,000未満、症例対照研究で100未満の場合、小さいと判定する。サンプルサイズの評価基準は信頼性評価の「2」点以下に相当する。)</p> <ul style="list-style-type: none"> ⊖⊖ 60%以上の文献で関連性指標の信頼性が低い ⊖ 関連性指標の信頼性が低い文献数が40%以上60%未満の時 ±0 関連性指標の信頼性が低い文献はあるが、その数が全体の40%未満の時
Up-grading	<p>(5) 関連性が強い (Magnitude: 関連性の強さを示す指標の推定値であるオッズ比またはリスク比の絶対値が2以上 (または0.5以下) の時に強いと判定)</p> <ul style="list-style-type: none"> ⊕⊕ 60%以上の文献で関連性が強い ⊕ 関連性が強い文献数が40%以上60%未満の時 ±0 関連性が強い文献はあるが、その数が全体の40%未満の時
Up-grading	<p>(6) 量反応関係がある (Dose response: 量反応関係が示された場合、エビデンスの総合評価を高める)。単に連続量として扱った解析での統計学的有意性、相関係数のみの提示では量反応関係の評価が難しいため量反応関係ありとはしない。</p> <ul style="list-style-type: none"> ⊕⊕ 60%以上の文献で量反応関係が示されている。 ⊕ 量反応関係を示した文献が40%以上60%未満の時 ±0 量反応関係を示した文献はあるが、その数が全体の40%未満の時
Up-grading	<p>(7) 制御された場合に関連性を強めるであろう残余交絡の存在 (Residual confounding)</p> <ul style="list-style-type: none"> ⊕⊕ 60%以上の文献で残余交絡がある。 ⊕ 残余交絡のある文献が40%以上60%未満の時 ±0 残余交絡のある文献はあるが、その数が全体の40%未満の時

Up- grading	(8) 異なる研究集団、研究デザインの研究間の結果の一致 (Consistency)
	⊕⊕ 60%以上の文献で結果の一致が示されている。
	⊕ 結果が一致する文献が40%以上60%未満の時
	±0 結果が一致する文献はあるが、その数が全体の40%未満の時

本研究では、上述の (1) から (8) を一覧表示し、エンドポイントのクラスターごとにBPAと当該エンドポイントのクラスターの関連性の信頼性を以下の如く判定した。

- 信頼度は高い (High certainty of evidence)
- 信頼度は中等度 (Moderate certainty of evidence)
- 信頼度は低い (Low certainty of evidence)

解析の結果、男性不妊、女性のIVF結果及び妊娠までの期間とBPAばく露との関連性は認められない (健康影響はない) という結果の証拠の信頼度は、中等度と判断された。一方、BPAばく露と精液の質または卵巣機能には関連性が認められたが、その証拠の信頼度は低いと判断された。

5 研究項目：最新の動物実験結果に基づく提言の取りまとめ

(1) 個別課題：BPA リスク評価への提言

(青山博昭 (残留農薬研究所)、井口泰泉 (横浜市立大学))

本研究で取り組んだ文献情報の収集とそれらの整理は、直接 BPA のリスク評価を目的としたものではなく、後に実施されるリスク評価に用いることができる信頼性の高いデータを使いやすい形式で提供すると共に、収集した文献情報がリスク評価のどの場面でどのように活用できるかを提言することを目的としたものであった。このため、BPA のリスク評価に際して常に問題となる低用量影響 (従来のリスク評価で NOAEL とされた 5 mg/kg bw/day 以下の用量域における何らかの生物学的変動) に着目して、それらが一望できる総括表 (別添資料表 5) を作成した。

今回の文献調査により、慢性毒性試験、発がん性試験、発生毒性試験、生殖毒性試験及び発達神経毒性試験を含む一般的な毒性試験で得られるエンドポイント (動物の体重、腫瘍性及び非腫瘍性病変の発生頻度、生殖能力に関わる各種エンドポイント、胎児発生に及ぼす影響及び発達神経毒性を評価するための検査指標等) に関しては、信頼性が担保されるすべての文献 (16 文献) を精査しても、一部の文献で報告された背景データの範囲内と考えられる変動を除き、5 mg/kg bw/day 以下の用量では BPA 投与の影響は検出されていないことが明らかとなった。したがって、今後のリスク評価に当たっても、これらの事実を念頭に議論を進めることができると考えられる。すなわち、既知の毒性影響については、これまでに得られている結論 (一般的な毒性指標を検討する限り、最も低い NOAEL は 5 mg/kg bw/day と判断する) を疑う余地はほとんどないと考えられる。

一方、通常の毒性試験では検査や観察が実施されない種々の観察項目については、比較的信頼性が高いと判断された文献においても、様々なエンドポイント (例えば、肝脂肪、血清アポプロテイン、nNOS 発現等) に 5 mg/kg bw/day 以下の用量域で対照群との有意な差が検出されていた。しかし、それらは、(1) 1 群当たりの観察動物数が、不足とまでは言えずとも必ずしも十分ではない実験で得られたデータ、(2) 毒性影響 (Adverse Outcome) であるとのコンセンサスが得られていないエンドポイントで観察された変化、(3) 過去に信頼できる解析がなく、正常範囲が不明なエンドポイントに観察された変動、のいずれかに該当するものがほとんどであった。

今後のリスク評価において、低用量域 (5 mg/kg bw/day 以下の投与群) で観察された毒性学的意義や再現性の有無が不明な変化をも含めて BPA のリスクを総合的に検討するのであれば、対照群との間

で有意な差が検出されたすべてのエンドポイントについて、(1) 実験系の信頼度を詳細に検討すること、(2) 同じエンドポイントを評価した他の文献から得られる情報に照らしてその再現性を吟味すること、(3) 作用機序を推定し、Adverse Outcome Pathway (AOP) 上のどこに位置するエンドポイントであるか、関連するエンドポイントにも同様の変化が生じているか、当該エンドポイントの正常範囲はどの程度であるか等を確認して、悪影響 (Adverse Outcome) であるか否かを正しく判定することが必要であると考えられる。また、作用機序が推定できないエンドポイントについては、そのような変化が BPA の投与以外の要因により引き起こされた可能性を十分に検討することが必要であると考えられる。

6 研究全体の成果、考察及び結論

本研究では、BPA のリスク評価に備えて関連文献を収集し、ヒト研究及び動物試験のそれぞれについて得られた情報の信頼度を評価するための統一的な基準を設定して、これらの文献から BPA のリスク評価に用いることができると考えられる信頼性の高い情報を抽出した。また、動物分野の研究から得られた情報については、それらの情報をリスク評価のどの場面でのどのような内容を用いるべきかを検討すると共に、低用量影響を含む BPA のリスクを評価する際に考慮すべき事項を整理した。さらに、疫学分野の文献についても、それらの信頼性を評価すると共に可能な限り情報を抽出して、現時点で得られる情報とそれらの信頼度を評価した。

動物実験分野の文献評価では、低用量域で観察されたとされる変化の多くについて、信頼性が高いと判断された文献から抽出したデータであっても毒性指標としての適切性や結果の再現性に疑問が残り、今後実施される BPA のリスク評価時に参照可能なデータは極めて限定されることが明らかとなった。また、それらの多くがやや特殊な (これまで毒性指標として用いられてこなかった) 指標に観察されたものであったため、これらの所見に関する背景データの蓄積もなく、作用機序を推定することも再現性を検討することも困難であった。毒性学的意義や再現性に疑問が残るこれらの変化については、リスク評価に先立って追加試験を実施することにより、BPA の低用量影響に関する知見をさらに深める必要があるかもしれない。また、米国で実施された CLARITY-BPA Study に関しては、大規模な動物実験で得られたサンプルの一部を用いて特定のエンドポイントを解析した研究者が自身の解析結果を個別に取りまとめた論文を公表し続けているので、今後もそれらの動向を注視する必要がある。

疫学分野の文献評価では、ヒトの生殖に及ぼす影響について、男性不妊、女性の IVF の結果及び妊娠までの期間といった指標に BPA ばく露との関連性は認められない (ヒトの生殖に対する明らかな影響はない) との文献情報が確認され、それらの証拠の信頼度は中等度であると判断された。一方、今回の解析により BPA ばく露と精液の質または卵巣機能には関連性が認められたが、その証拠の信頼度は低いと判断された。これらの情報は、後に実施されるリスク評価の際に、動物実験の結果を補完することができると考えられる。しかし、疫学研究ではばく露量を正確に定量することができないため、疫学分野の研究成果から低用量影響評価に有用な情報を抽出することは困難であると考えられる。今後は、他のエンドポイントに関するクラスターについても証拠の信頼度のランク付けを行うと共に、それらのクラスターごとに証拠の信頼度のランク付けに資するようなメタアナリシス (研究間の異質性や出版バイアスの評価、統合効果量の推定などを含む) を実施することが望まれる。また、分析 (メタアナリシス) の精度を高めるためには、疫学研究におけるばく露やエンドポイントの標準化が必要であると考えられる。

Ⅲ 本研究を基にした論文等

- 1 本研究を基にした論文と掲載された雑誌名のリスト
なし

- 2 本研究を基にした学会発表の実績
なし

- 3 特許権等の出願・申請等の状況
なし

- 4 プログラムの著作物及びデータベースの著作物
なし

- 5 その他（各種受賞、プレスリリース等）
なし

IV 研究開始時に申告した達成目標及び研究全体の自己評価

1 達成目標の自己評価

達成目標	評価結果	自己評価コメント
(1) BPA のリスク評価に向けて、有用性を評価すべき最大約800文献を対象として可能な限りこれらの文献を収集し、内容に応じて体内動態調査、毒性評価、ばく露量評価、及び疫学調査の4分野に整理する。	5	ほぼ予定通りに文献を収集し終え、これらを整理してデータベース化することができた。
(2) 動物実験分野及び疫学分野の文献について、それらの信頼度を評価する適切な基準を設定し、BPA のリスク評価に用いることができる信頼性の高い文献を選抜する。	5	2分野の文献を対象に、それらの信頼性を評価するための基準を設定し、信頼性の高い文献を抽出した。
(3) BPA のリスク評価に資するよう、選抜された文献からデータを抽出してそれらをエンドポイント別に整理し、これらの文献の適切な参照範囲とその用途を提言する。	3	信頼性が高いと判断された実験動物分野の文献からデータを抽出し、エンドポイント毎に整理した。また、疫学分野の文献については、生殖関連のエンドポイントについて情報の信頼度を評価した。しかし、BPA の低用量影響に関する懸念をすべて解消することはできず、リスク評価に当たって検討すべき事項を提言するに止まった。

注) 評価結果欄は「5」を最高点、「1」を最低点として5段階で自己採点。

2 研究全体の自己評価

項目	評価結果	自己評価コメント
(1) 研究目標の達成度	4	BPA の低用量影響に関する懸念を払拭することはできなかったが、議論の方向性に関する提言を示すことはできたと考える。
(2) 研究成果の有用性	4	今後実施されるBPAのリスク評価に当たり、参照すべき信頼性の高いデータを整理して提示することができた。
総合コメント これまでに公表された文献を整理し、信頼性が高いと考えられるデータを抽出して整理することができた。また、リスク評価に当たって考慮すべき点を提示した。		

注) 評価結果欄は、「5」を最高点、「1」を最低点として5段階で自己採点。

この報告書は、食品安全委員会の委託研究事業の成果について取りまとめたものです。

本報告書で述べられている見解及び結論は研究者個人のものであり、食品安全委員会としての見解を示すものではありません。全ての権利は、食品安全委員会に帰属します。

表 1. 動物実験分野の文献の信頼性評価基準

分類	評価項目	評価法	評価内容及び基準	備考	EFSA の Risk of Bias Tool との対応
(1) 研究体制	研究目的	○/×	リスク評価に用いることを前提にした危害分析 (Hazard identification) を目的としたものか、メカニズム解析等を目的としたものかの区分 (メカニズム解析のみを目的とした論文は除く)	必須	— (Relevancy に関係)
	実験規模	1~4	試験群の構成 (対照群/陽性対照群を含む) 及び 1 群当りの動物数は適切か 各群に十分な数の動物がほぼ同数割り振ら得ている : 4 群ごとの動物数は必ずしも一定ではないが、必要な動物数は確保されている : 3 各群にほぼ同数の動物が割り振られているものの、その数はやや少ない : 2 各群に割り振られた動物数はまちまちで、数も少ない : 1		(8) Other sources of bias
	動物飼育環境	1~4	動物の飼育環境は適切か SPF 施設で、十分な配慮の下に均一な飼育条件が担保されている : 4 Conventional な飼育施設で、均一な飼育条件が担保されている : 3 一般的な飼育施設に動物を収容しているが、群間で飼育条件が異なる : 2 動物の飼育環境に関して、必要な配慮に欠ける : 1		(6) Detection (outcome assessment)
	動物愛護	○/×	AAALAC International (The Association for Assessment and Accreditation of Laboratory Animal Care International) または研究実施国が定める基準を満たしているか。	参考情報	—

	ガイドライン/GLP 準拠の有無	○/×	標準的なガイドライン/GLP に準拠していれば、データの採取や取り扱いについて、一定の信頼を与えることができる	参考情報	(1) Selection (randomised) 等
	盲検法の採用	○/×	動物の観察者に投与用量を伏せて検査を実施させるべき指標（行動検査等）について、適切な対応が取られているか	参考情報	(2) Selection (concealed) (6) Detection
(2) 材料及び 方法	BPA に関する記載	1~4	基本情報（入手先、ロット、純度、溶媒及びその濃度等）が適切に記載されているか すべての情報が正しく記載されている：4 一部の情報が欠けるものの、必要最低限の情報は記載されている：3 記載が漠然としており、やや曖昧である：2 必要不可欠な情報が欠落している：1		(5) Detection (exposure characterisation)
	BPA のばく露	1~4	適切な投与経路（経口投与）と用量（3 用量以上）が選定されているか 3 用量以上の経口投与群が適切に設定されている：4 投与経路や投与群数に問題はないが、用量設定が必ずしも適切でない：3 投与経路に問題はないが、用量群数が不足している：2 投与経路または用量設定が適切でない：1	投与経路については一次評価で確認済み	— (Relevancy に関係)
	実験動物の選択	1~4	実験目的に合った系統またはストックが選択されているか 遺伝的に均一な近交系のラットまたはマウスが使用されている：4 一般的なアウトブリード・ストックのラットまたはマウスが使用されている：3 遺伝的にほぼ均一なラット・マウス以外の種が使用されている：2 遺伝的統御がなされていない実験動物を用いている：1		(6) Detection (outcome assessment)

	動物の飼育条件	1~4	動物の飼育条件は群間で均一か 1 ケージ当たりの動物数は適切であり、投与または検査のストレスも制御されている：4 1 ケージ当たりの動物数は適切だが、投与または検査時の過剰なストレスが疑われる：3 一部の群で飼育条件または動物に与えるストレスが異なる：2 必要な配慮に欠ける（明確な記載がなく、判断不能）：1		(3) Performance
	基礎飼料の成分	1~4	基礎飼料に含まれる栄養成分に含まれる植物エストロゲンの扱いは妥当か 植物エストロゲン・フリーの飼料（AIN-93G 等）を用いている：4 植物エストロゲン・フリーではないが、実験結果に及ぼす影響が考慮されている：3 一般的な基礎飼料が使用されている：2 用いた基礎飼料に関する記載がない：1		(5) Detection (exposure characterisation)
	飲料水、溶媒及び飼育器材の管理	1~4	BPA または BPA と同様の生体反応を引き起こす可能性のある物質による汚染がないか 適切な方法（化学分析等）で汚染がないことを確認している：4 分析は実施していないが、汚染源と推測される材料の使用は避けられている：3 通常の飼育方法が用いられており、汚染はないと推測される：2 汚染の有無について判断できない、または、汚染の恐れがある：1		(5) Detection (exposure characterisation)
	統計学的手法	1~4	得られたデータが適切な統計学的解析結果に基づいて科学的に評価されているか すべての指標について、標本単位及び統計学的解析手法が適切と判断される：4 一部の指標について、標本単位または統計学的解析手法に妥当性を欠く：3 標本単位または統計学的手法に妥当性を欠く指標が散見される：2 不適切な解析が実施されている：1		(8) Other sources of bias
(3) 結果	データの取り扱い	1~4	特定の動物（例えば片性の哺育児のみ）を意図的に選抜したり除外したりしていないか すべての動物を評価対象にしている：4 一部の動物が評価から除外されているが、合理的な理由が明記されている：3 特段の理由なく一部の動物が評価から除外されている：2 多くの検査指標について、評価対象動物が明らかでない：1		(4) Attrition

観察指標及びそれらの取り扱い	1~4	科学的及び生物学的に妥当な観察指標が設定され、それらが欠落なく報告されているか 妥当な観察指標が設定され、それらが欠落なくすべて報告されている：4 妥当な観察指標が設定されているものの、結果の一部が報告されていない：3 すべての結果が報告されているものの、一部の指標は妥当性を欠く：2 観察指標に妥当性を欠く：1		(7) Selective reporting
標本単位	1~4	標本単位（生殖・発生毒性学の分野では、通常、腹を標本単位とする）は適切か すべてのデータについて、標本単位は適切である：4 一部のデータについて、標本単位が必ずしも適切ではない*：3 不適切な標本単位が散見される：2 標本単位が明らかでない：1		(8) Other sources of bias
データの表示	1~4	定量的なデータについて、各群の平均値と標準偏差（または標準誤差）が明示されているか 各群の平均値と標準偏差（または標準誤差）が適切に表示されている：4 平均値は表示されているものの、誤差範囲が不明瞭である：3 何らかの形で測定値が示されているものの、平均値が明示されていない：2 一部のデータが欠落している：1		— (Qualityに 関係)
図（写真）	1~4	病理組織学的検査や分子生物学的解析の結果等が、十分な品質の図として示されているか すべての図の品質が十分に高い：4 一部に不明瞭な図が含まれるものの、データの信頼性に問題はない：3 品質の低い図が含まれるため、一部のデータについては信頼性に欠ける：2 データの信頼性を保証するために必要不可欠な図が示されていない：1		— (Qualityに 関係)

*:胎児または哺育児に観察された所見については、各群の総出現頻度を用いた評価も可能な場合がある。

表2 信頼性が高いと評価された動物実験分野の文献一覧

データベース上の文献番号	著者	論文表題	掲載誌
11	Ronn M <i>et al.</i>	Bisphenol A exposure increases liver fat in juvenile fructose-fed Fischer 344 rats.	<i>Toxicology</i> , 303:125-132, 2013.
17	Bauer SM <i>et al.</i>	The effects of maternal exposure to bisphenol A on allergic lung inflammation into adulthood.	<i>Toxicol. Sci.</i> , 130:82-93, 2012.
80	Martini M <i>et al.</i>	Effects of perinatal administration of bisphenol A on the neuronal nitric oxide synthase expressing system in the hypothalamus and limbic system of CD1 mice.	<i>J. Neuroendocrinol.</i> , 22:1004-1012, 2010.
119	Ryan BC <i>et al.</i>	In utero and lactational exposure to bisphenol A, in contrast to ethinyl estradiol, does not alter sexually dimorphic behavior, puberty, fertility, and anatomy of female LE rats.	<i>Toxicol. Sci.</i> , 114:133-148, 2010.
132	Tyl RW <i>et al.</i>	Two-generation reproductive toxicity study of dietary bisphenol A in CD-1 (Swiss) mice.	<i>Toxicol. Sci.</i> , 104:362-384, 2008.
133	Tyl RW <i>et al.</i>	Three-generation reproductive toxicity study of dietary bisphenol A in CD Sprague-Dawley rats.	<i>Toxicol. Sci.</i> , 68:121-146, 2002.
164	Kendzierski JA <i>et al.</i>	Strain specific induction of pyometra and differences in immune responsiveness in mice exposed to 17 α -ethinyl estradiol or the endocrine disrupting chemical bisphenol A.	<i>Reprod. Toxicol.</i> , 34:22-30, 2012.
171	Stump DG <i>et al.</i>	Developmental neurotoxicity study of dietary bisphenol A in Sprague-Dawley rats.	<i>Toxicol. Sci.</i> , 115:167-182, 2010.
192	Delclos KB <i>et al.</i>	Toxicity evaluation of bisphenol A administered by gavage to Sprague Dawley rats from gestation day 6 through postnatal day 90.	<i>Toxicol. Sci.</i> , 139:174-197, 2014.
608	Folia M <i>et al.</i>	Oral homeostasis disruption by medical plasticizer component bisphenol A in adult male rats.	<i>Laryngoscope</i> , 123:1405-1410, 2013.

696	Lejonklou MH <i>et al.</i>	Low-dose developmental exposure to bisphenol A alters the femoral bone geometry in Wistar rats.	<i>Chemosphere</i> , 164:339-346, 2016.
698	Leung YK <i>et al.</i>	Gestational high-fat diet and bisphenol A exposure heightens mammary cancer risk.	<i>Endocrine-Related Cancer</i> , 24:365-378, 2017.
729	Mandrup K <i>et al.</i>	Low-dose effects of bisphenol A on mammary gland development in rats.	<i>Andrology</i> , 4:673-683, 2016.
750	National Toxicology Project	Draft NTP research report on the CLARITY-BPA Core Study: A perinatal and chronic extended-dose-range study of bisphenol A in rats.	<i>NTP Research Report 9</i> , 2018.
776	Patel S <i>et al.</i>	Bisphenol A exposure, ovarian follicle numbers, and female sex steroid hormone levels: Results from a CLARITY-BPA Study.	<i>Endocrinology</i> , 158:1727-1738, 2017.
794	Rebuli ME <i>et al.</i>	Impact of low-dose oral exposure to bisphenol A (BPA) on juvenile and adult rat exploratory and anxiety behavior: A CLARITY-BPA Consortium Study.	<i>Toxicol. Sci.</i> , 148:341-354, 2015.

表 3. リスク評価に適さないと判断された動物実験分野の文献から得られたデータ

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.002 0.0025	0.005	0.001	0.02 0.025 0.03	0.04 0.05	0.1	0.2 0.25	0.3	0.49 0.5	2.5	5	12.5	30	50	150 500
11 F344 Rat 5-15 週齢、剖 検 15 週齢 3	F0 ♀ 体重		○			○				○						
	F0 ♀ 皮下脂肪量		○			○				○						
	F0 ♀ 肝脂肪量 ↑		○			●				●						
	F0 ♀ 血清アポリポタンパク ↑		○			●				●						

○：対照群との間に統計学的有意差はない。 ●：対照群との間に統計学的有意差がある。

注：BPA 投与による影響と考えられるが、果糖との併用投与によるものであること、有意差のあるエンドポイントの毒性学的意義は低いことから、BPA の毒性評価には適さないと考えられる。

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.002 0.0025	0.005	0.001	0.02 0.025 0.03	0.04 0.05	0.1	0.2 0.25	0.3	0.49 0.5	2.5	5	12.5	30	50	150 500
17 C57BL/6J Mouse GD6-PND21 4	F1 ♀ 肺炎症細胞浸潤 ↑									○		○			●	●
	F1 ♂ 肺炎症細胞浸潤 ↑									○		○			●	●
	F1 ♀ 肺炎症細胞浸潤 ↓									○		○			●	○
	F1 ♂ 肺炎症細胞浸潤									○		○			○	○
	F1 ♀ 気道好酸球浸潤 ↓									●		●			●	○
	F1 ♀ 肺好酸球浸潤 ↓									●		●			●	○

○：対照群との間に統計学的有意差はない。 ●：対照群との間に統計学的有意差がある。

注：気道過敏性に変化は認められず、これらの変化の毒性学的意義は低いことから、BPA の毒性評価には適さないと考えられる。

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.002 0.0025	0.005	0.01	0.02 0.025 0.03	0.04 0.05	0.1	0.2 0.25	0.3	0.49 0.5	2.5	5	12.5	30	50	150 500
80 CD-1 Mouse GD11-PD8 剖検生後2カ月 3	F1♀体重↑			○	○	○										
	F1♂体重↑			○	○	○										
	F1♀nNOS (MPOM) 発現↑			●	●	○										
	F1♂nNOS (MPOM) 発現			○	○	○										
	F1♀nNOS (BSTmv) 発現			○	○	○										
	F1♂nNOS (BSTmv) 発現↓			○	●	○										

○：対照群との間に統計学的有意差はない。 ●：対照群との間に統計学的有意差がある。

注：BPA 投与による影響の可能性はあるが、いずれも毒性的変化ではない（現時点では毒性学的意義が明らかでない）ことから、BPA の毒性評価には適さないと判断される。

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.002 0.0025	0.005	0.01	0.02 0.025 0.03	0.04 0.05	0.1	0.2 0.25	0.3	0.49 0.5	2.5	5	12.5	30	50	150 500
119 LE Rat/ GD7-PND18 2-3	F1♀AGD PND2				○			○								
	F1♀体重 PND2	○			○			○								
	F1♀体重離乳日				○			○								
	F1♀膣開口日	○			○			○								
	F1♀膣開口日体重	○			○			○								
	F2 産児数				○			○								
	F1♀妊娠率				○			○								
	F1♀サッカリン嗜好性	○			○			○								
	F1♀自発運動性	○			○			○								
	F1♀発情反射行動				○			○								
	外性器発達異常				○			○								
	尿道口膣口間距離 (10 か月)	○			○			○								

○：対照群との間に統計学的有意差はない。

注： いずれの観察項目にも有意差は得られていない。

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.002	0.005	0.0	0.02	0.04	0.1	0.2	0.3	0.49	2.5	5	12.5	30	50	150
164 C57BL/J CD-1 Mouse 6-23 週 3	F0 ♀BL 体重				○				○					○		
	F0 ♀CD1 体重				○				○					○		
	F0 ♀BL 受胎率			1	○				○					○		
	F0 ♀CD1 受胎率				○				○					○		
	F0 ♀BL 子宮重量				○				○					○		
	F0 ♀CD1 子宮重量				○				○					○		
	F0 ♀BL 組織学的異常 ↑				○				●					○		
	F0 ♀CD1 組織学的異常				○				○					○		
マクロファージ のみ 1	F0 ♀BL 子宮マクロファージ浸潤 ↑								●							
	F0 ♀CD-1 子宮マクロファージ浸潤								○							

○：対照群との間に統計学的有意差はない。 ●：対照群との間に統計学的有意差がある。

注：1例に観察された子宮蓄膿症はBPA投与により誘発された可能性を否定できないが、数理モデルによりエストロゲン感受性の系統差を示すものであり、BPAの毒性評価には適さないと判断される。

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.002 0.0025	0.005	0.01	0.02 0.025 0.03	0.04 0.05	0.1	0.2 0.25	0.3	0.49 0.5	2.5	5	12.5	30	50	150 500
171 SD Rat/ GD0-PND21 5	F0♀体重 GD3-GD20 ↓			○			○					○			●	●
	F0♀体重 LD1-7, 14 ↓			○			○					○			○	●
	F0♀体重 LD11, 14 ↓			○			○					○			●	●
	F1♂体重 PND4-11 ↑			○			○					●			○	○
	F1♂体重 PND14, 17 ↓			○			○					○			●	●
	F1♂体重 PND21 ↑ ↓			● ↑			○					● ↑			○	● ↓
	F1♂♀体重 PND28-72			○			○					○			○	○
	F1♀行動試験 PND13-62			○			○					○			○	○
	F1♂♀脳形態計測 PND21-72			○			○					○			○	○

○：対照群との間に統計学的有意差はない。 ●：対照群との間に統計学的有意差がある。 ↑統計学的に有意に増加。 ↓統計学的に有意に低下。

注：低用量での体重の変動は、用量相関傾向が明確ではなく、一貫した時間的傾向がみられないことから、BPA 投与によるものとは考えられない。

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.002 0.0025	0.005	0.001	0.02 0.025 0.03	0.04 0.05	0.1	0.2 0.25	0.3	0.49 0.5	2.5	5	12.5	30	50	150 500
608 Wistar Rat 12 週齢 5	F0♂飲水量 ↓		○			○						○	●			
	F0♂食塩嗜好性 ↑		○			○				○		●	●			
	F0♂顎下腺腺房減少 ↓		○			○				○		○	●			

○：対照群との間に統計学的有意差はない。 ●：対照群との間に統計学的有意差がある。

注：BPA 投与による影響の可能性はあるが、いずれも毒性学的に意味のある変化ではなく、BPA の毒性評価には適さないと判断される。

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.002 0.0025	0.005	0.001	0.02 0.025 0.03	0.04 0.05	0.1	0.2 0.25	0.3	0.49 0.5	2.5	5	12.5	30	50	150 500
696 Wistar Rat GD7-PD22 4	F1♀体重 (3 カ月)				○				○						○	
	F0♂体重 (3 カ月)				○				○						○	
	F1♀大腿骨長 (3 カ月) ↑				●				○						●	○

○：対照群との間に統計学的有意差はない。 ●：対照群との間に統計学的有意差がある。

注：BPA 投与による影響の可能性はあるが、極めてわずかな変化であり、GLP への準拠、盲検法の有無についての記載もないことから、BPA の毒性評価には適さないと判断される。

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.002 0.0025	0.005	0.001	0.02 0.025 0.03	0.04 0.05	0.1	0.2 0.25	0.3	0.49 0.5	2.5	5	12.5	30	50	150 500
698 SD Rat GD9-PD0 4	F1♀体重 PD2-50	○			○				○						○	
	F1♀ホルモンレベル PD21/50	○			○				○						○	
	F1♀乳腺末梢芽突起数 PND21 ↑	○			●				○						○	
	F1♀乳腺腫瘍 PND140 ↑	○			●				○						○	

○：対照群との間に統計学的有意差はない。 ●：対照群との間に統計学的有意差がある。

注：組織学的変化についてはいずれも BPA 投与の影響と推測されたが、高脂肪食の併用及び DMBA 誘発腫瘍の変化であることから、BPA の毒性評価には適さないと判断される。

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.002 0.0025	0.005	0.001	0.02 0.025 0.03	0.04 0.05	0.1	0.2 0.25	0.3	0.49 0.5	2.5	5	12.5	30	50	150 500
729 Wistar Rat GD7-PND22 4	F1♀乳腺の長さ PND22				○				○			○			○	
	F1♂乳腺の長さ PND22				●				○			○			○	
	F1♀乳腺組織変化 intraductal hyperplasia PND400 ↑				○				●			○			○	
	F1♂乳腺組織変化 PND100				○				○			○			○	

○：対照群との間に統計学的有意差はない。 ●：対照群との間に統計学的有意差がある。

注：いずれも投与による影響の可能性があるとしたが、病変の背景データが得られておらず、GLP への準拠や盲検法の採用についても記載がないため、BPA の毒性評価には適さないと判断される。

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.002 0.0025	0.005	0.001	0.02 0.025 0.03	0.04 0.05	0.1	0.2 0.25	0.3	0.49 0.5	2.5	5	12.5	30	50	150 500
794 SD Rat GD6-出産 PND1-PND21 3	F1♂行動試験 PND25	○			○						○					
	F1♀行動試験 PND25 迷路、オープンフィールド等	○			○						○					
	F1♂♀行動試験 PND26-27, 97-111, 111-125	○			○						○					

○：対照群との間に統計学的有意差はない。

注： いずれの観察項目にも有意差は得られていない。

表 4. 動物実験分野の文献から抽出されたリスク評価に有用と考えられるデータ

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.001	0.0025	0.008	0.02 0.025	0.08	0.25 0.26	0.3	0.84	2.5 2.7	4.39 5	25	45.9 50	100	300	500 514
132 ICR Mouse F0 世代交 配前 8 週 - F2 世代 PND21 混餌 6	F1♂♀体重↓		○		○		○				○		○			●
	F1♂♀肝臓重量↑		○		○		○				○		○			●
	F1♀左腎臓重量↑		○		○		○				○		○			●
	F2♂Weanlings 脾臓重量↓		○		○		○				○		○			●
	F2♂Weanlings 精巣重量↓		○		○		○				○		○			●
	F2♂包皮分離日↑		○		○		○				○		○			●
	F1, F2♀脾臓重量↓		○		○		○				○		○			●

○ : 対照群との間に統計学的有意差はない。

● : 対照群との間に統計学的有意差がある。

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.001	0.0025	0.008	0.02 0.025	0.08	0.25 0.26	0.3	0.84	2.5 2.7	4.39 5	25	45.9 50	100	300	500 514
133 SD Rat F0 世代交 配前 10 週 - F3 世代 PND21 混餌 6	♂♀体重↓	○			○			○			○		●			●
	着床数・産児数↓	○			○			○			○		○			●
	F1♀膈開口日齢↑	○			○			○			○		○			●
	F1♂包皮分離日齢↑	○			○			○			○		○			●
	♀尿管退縮率/ 肝臓炎症↑	○			○			○			○		○			●
	♂精囊・凝固重量体重比 ↑	○			○			○			○		○			
	♀卵巣重量体重比↓	○			○			○			○		○			●
192 SD Rat GD6-PND90 強制経口 9	体重♂↓		○	○	○	○	○		○	○				○	●	
	♀↓		○	○	○	○	○		○	○				●	●	
	F1♂T3、T4、TSH 濃度↑		○	○	○	○	○		○	○				●	●	
	黄体数↓		○	○	○	○	○		○	○				○	●	

○：対照群との間に統計学的有意差はない。

●：対照群との間に統計学的有意差がある。

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.001	0.0025	0.008	0.02 0.025	0.08	0.25 0.26	0.3	0.84	2.5 2.7	4.39 5	25	45.9 50	100	300	500 514
750 SD Rat F0 雌 GD6- 出産, PND1-最長 2年間 強制経口 5	<u>1年齢 (n=20-26) :</u>															
	子宮 lesion ↑		○		○		○			○		●				
	卵巣 lesion ↑		○		○		○			○		●				
	♀肝臓 lesion ↑		●		○		○			○		●				
	♂肝臓 lesion ↑		○		○		●			●		○				
	精巣上体リンパ球浸潤 ↑		○		○		○			○		●				
	<u>2年齢 (n=46-50) :</u>															
	膾上皮過形成 ↑		○		●		○			●		○				
	♀腎臓 lesion ↑		●		○		○			○		○				
	♀肝臓 lesion ↑		●		○		○			○		●				
	甲状腺シスト ↑		○		○		●			●		○				
	前立腺 lesion ↑		●		○		○			○		○				
	下垂体 lesion ↑		○		○		○			○		●				
♂甲状腺 C 細胞増殖 ↑		○		○		○			○		●					
♂副甲状腺細胞過剰増殖 ↑		○		○		●			○		○					
♂腎臓肥大 ↑		○		○		●			○		○					

○ : 対照群との間に統計学的有意差はない。

● : 対照群との間に統計学的有意差がある。

文献番号 動物種 投与期間 投与群数	エンドポイント	投与用量 (mg/kg/day)														
		0.001	0.0025	0.008	0.02 0.025	0.08	0.25 0.26	0.3	0.84	2.5 2.7	4.39 5	25	45.9 50	100	300	500 514
776 SD Rat F0 雌 GD6- 出産, PND1-最長 1年間 強制経口 5	<u>PND21 (n=8-10) :</u> 原始濾胞数 ↓		○		○		●			○		○				
	一次濾胞数 ↓		●		○		○			○		○				
	前胎状濾胞数 ↓		●		○		○			○		○				
	総正常濾胞数 ↓		●		○		●			○		○				
	<u>1 年 齢 (n=8-10) :</u> E2 濃 度 ↓		○		○		○			●		●				

観察には、文献 750 で実施された試験で得られた F1 世代の動物（各群 46-50 腹の同腹児）の一部を無作為に抽出して用いた。

○：対照群との間に統計学的有意差はない。

●：対照群との間に統計学的有意差がある。

表 5. エンドポイント別に整理した今後のリスク評価に有用と考えられるデータ

文献番号 動物種	投与期間 エンドポイント：体重	投与用量 (mg/kg/day)														
		0.001	0.0025	0.008	0.02 0.025	0.08	0.25 0.26	0.3	0.84	2.5 2.7	4.39 5	25	45.9 50	100	300	500 514
132 : ICR Mouse	交配前 8w-PND21 F1♂♀		○		○		○				○		○			↓
133: SD Rat	交配前 10w-PND21 F1♂♀	○			○			○			○		↓			↓
192: SD Rat	GD6-出産 : F1♂ F1♀		○	○	○	○	○		○	○			○	↓	↓	

○ : 対照群との間に統計学的有意差はない。

↓ 統計学的に有意に低下。

文献番号 動物種	投与期間 エンドポイント：器官重量	投与用量 (mg/kg/day)														
		0.001	0.0025	0.008	0.02 0.025	0.08	0.25 0.26	0.3	0.84	2.5 2.7	4.39 5	25	45.9 50	100	300	500 514
132 : ICR Mouse	交配前 8w-PND21 F1♂♀肝臓重量 F1♀左腎臓重量 F1/F2♀脾臓重量 F2 ♂Weanlings 脾臓重量 精巣重量		○		○		○				○		○			↑ ↑ ↓ ↓ ↓
133: SD Rat	交配前 10w-PND21 F1♂精囊・凝固腺重量体重比 F1♀卵巣重量体重比	○			○			○			○		○			↑ ↓

○ : 対照群との間に統計学的有意差はない。

↑ 統計学的に有意に増加。 ↓ 統計学的に有意に低下。

文献番号 動物種	投与期間 エンドポイント:生殖関連指標	投与用量 (mg/kg/day)														
		0.001	0.0025	0.008	0.02 0.025	0.08	0.25 0.26	0.3	0.84	2.5 2.7	4.39 5	25	45.9 50	100	300	500 514
132 : ICR Mouse	交配前 8w-PND21 F2♂包皮分離日		○		○		○					○	○			↑
133 : SD Rat	交配前 10w-PND21 F1♂包皮分離日 F1♀膻開口日 着床数・産児数	○			○			○			○	○				↑ ↑ ↓
192 : SD Rat	GD6-出産 : ♀ 黄体数		○	○	○	○	○		○	○			○	↓		

○ : 対照群との間に統計学的有意差はない。

↑ 統計学的に有意に増加。 ↓ 統計学的に有意に低下。

文献番号 動物種	投与期間 エンドポイント:組織異常	投与用量 (mg/kg/day)															
		0.001	0.0025	0.008	0.02 0.025	0.08	0.25 0.26	0.3	0.84	2.5 2.7	4.39 5	25	45.9 50	100	300	500 514	
133: SD Rat	♀尿細管退縮率/肝臓炎症	○			○			○			○		○			↑	
750: SD Rat	GD6-出産 <u>1年齢 (F1):</u>																
	子宮 lesion		○		○			○			○		↑				
	卵巣 lesion		○		○			○			○		↑				
	♀肝臓 lesion		↑		○			○			○		↑				
	♂肝臓 lesion		○		○			↑			↑		○				
	精巣上体リンパ球浸潤		○		○			○			○		↑				
	<u>2年齢 (F1):</u>																
	膺上皮過形成		○		↑			○			↑		○				
	♀腎臓 lesion		↑		○			○			○		○				
	♀肝臓 lesion		↑		○			○			○		↑				
	甲状腺シスト		○		○			↑			↑		○				
	前立腺 lesion		↑		○			○			○		○				
	下垂体 lesion		○		○			○			○		↑				
	♂甲状腺C細胞増殖		○		○			○			↑		○				
♂副甲状腺細胞過剰増殖		○		↑			○			○		○					
♂腎臓肥大		○		↑			○			○		○					

○: 対照群との間に統計学的有意差はない。

↑ 統計学的に有意に増加。

文献番号 動物種/ 投与期間	投与期間 エンドポイント: <u>卵巣組織/ ホルモン濃度</u>	投与用量 (mg/kg/day)														
		0.001	0.0025	0.008	0.02 0.025	0.08	0.25 0.26	0.3	0.84	2.5 2.7	4.39 5	25	45.9 50	100	300	500 514
776: SD Rat	<u>GD6-出産</u>															
	<u>PND21:</u>															
	原始濾胞数		○		○		↓			○		○				
	一次濾胞数		↓		○		○			○		○				
	前胞状濾胞数		↓		○		○			○		○				
総正常濾胞数		↓		○		↓			○		○					
1 年 齢 (F1):																
	E2 濃度		○		○		○			↓		↓				
192: SD Rat	<u>GD6-出産</u> F1♂T3、T4、TSH 濃度		○	○	○	○	○		○	○				↑	↑	

○ : 対照群との間に統計学的有意差はない。

↑ 統計学的に有意に増加。 ↓ 統計学的に有意に低下。

表 6. 疫学分野の文献の信頼性評価基準

分類		評価項目	評価法	評価内容及び基準		EFSA の Risk of Bias Tool との対応
				コホート研究	症例対照研究	
材料及び方法	1	集団の比較可能性	1~4	<p>ばく露の有無にかかわらず研究対象者の特徴は類似しているという直接的なエビデンスがある。</p> <ol style="list-style-type: none"> ばく露有無が異なる集団の比較となっていない 研究対象者全体で同じ方法を用いてばく露有無の判定が行われている ばく露の有無によらず、選択・除外基準は異なる ばく露群間で年齢や健康状態は類似している（曝露の有無別でベースライン特徴に統計学的な差がない） ばく露有無の測定は同時期に全て実施され、ばく露の有無により参加・同意率は異なる。 <p>1~5 の観点から 適切である:4 概ね適切である:3 やや不適切である:2 不適切である:1</p>	<p>症例と対照は類似しているという直接的なエビデンスがある</p> <ol style="list-style-type: none"> 対照は症例と同じ母集団からの抽出である 症例・非症例（対照）の判定は同じ方法を用いてなされ、症例の選択基準は対照の選択基準と矛盾しない 症例と対照の年齢、性別、民族などは類似している 症例と対照は同時に研究に組み入れた 対照は研究対象疾患の既往歴を有しないことが記述されている。 <p>1~5 の観点から 適切である:4 概ね適切である:3 やや不適切である:2 不適切である:1</p>	(1) Selection
	2	対象者数または症例数	1~4	<p>対象者数</p> <p>10,000 以上:4 5,000-9,999:3 1,000-4,999:2 1,000 未満:1</p>	<p>症例数</p> <p>500 以上:4 100-499:3 50-99:2 50 未満:1</p>	(7) Other sources of bias

3	エンドポイント脱落による除外	1~4	<p>追跡不能による対象者の除外に関して明らかな問題がない。</p> <ol style="list-style-type: none"> 1. 当該事実の有無および理由の明記あり 2. 妥当な範囲の対象者数減少である 3. 追跡不能理由がエンドポイントに無関係 4. 追跡不能理由がばく露によって異なる 5. 追跡不能者の特徴も記述され、追跡不能により除外されなかった対象者との違いは統計学的に有意ではない。 <p>1~5の観点から 適切である:4 概ね適切である:3 やや不適切である:2 不適切である:1</p>	<p>対象者（症例及び対照）の解析からの除外がもしある場合は、以下のような明らかな問題がない。</p> <ol style="list-style-type: none"> 1. 除外されていて理由の記載がない 2. 除外の理由がエンドポイント定義に関連している可能性がある。その状況として、例えば、症例と対照で除外される者の割合が異なる。あるいは症例と対照で除外の理由が異なる。 <p>1~2の観点から 適切である:4 概ね適切である:3 やや不適切である:2 不適切である:1</p>	(2) Attrition
4	ばく露欠測による除外	1~4	<p>除外基準が明確に述べられており、対象者がばく露欠測の理由により解析から除外される場合、その理由が明記されている。ばく露欠測がある場合、対象者全体に均等に起きていると想定され、欠測のある集団の特徴が記述されている。</p> <p>明記されていて、問題ない（欠測のある変数の解析上の取り扱いも含む）:4 一部記載がある:3 記載がない:2 除外方法に問題がある:1</p>		(3) Detection (exposure characterisation)
5	ばく露測定方法	1~4	<p>全対象者に対して同一方法で同一時期にばく露の測定が実施されたことが明らかで、かつその方法は生体へのばく露評価として確立した直接的なもの（環境中・生体試料中の化学物質の測定、例：1年間に3回の24時間蓄尿による評価）であるかまたは、その方法の妥当性が、より確立した方法を参照基準として確認されている場合：4</p> <p>上記条件を満たしていない場合で、全対象者に対して同一方法で同一時期にばく露の測定が実施されたと想定されるか、多少の差異があっても、そのことにより偏った研究結果が得られていないと考えられる場合、あるいは、測定方法が上記の確立した方法に対する妥当性が確認された間接的なばく露評価法である質問紙や作業環境測定士による作業環境測定結果を用いている場合：3</p> <p>ばく露群と非ばく露群（コホート研究の場合）に対し、あるいは症例と対照（症例対照研究の場合）に対し、異なる方法、異なった時期にばく露が測定されたと想定されるか十分な情報が記載されていない場合、測定方法は直接的なものであるが確立していない方法で</p>		(3) Detection (exposure characterisation)

			<p>あつたり、妥当性の確認が十分でない間接的なもの（職業、自己申告）の場合、あるいはばく露測定に関する記載が十分でない場合：2</p> <p>ばく露群と非ばく露群（コホート研究の場合）における、あるいは症例と対照（症例対照研究の場合）におけるばく露測定方法や時期が明らかに異なる場合やばく露測定方法の妥当性が明らかに問題の場合：1</p>	
6	ばく露指標の妥当性および対象集団のばく露の変動範囲	1~4	<p>ばく露指標は、健康影響が想定される月や年単位でのばく露評価となっているか。研究対象となる母集団で予想されるばく露の範囲を研究対象集団が有しているか。コホート研究においては、ばく露の個人間差が十分に検出できる対象を選んでいるか。症例対照研究においては検出されたばく露の差が生物学的に意味のある範囲となっているか。</p> <p>例えば、小児期の精神発達をエンドポイントにした研究で、妊娠中の母のBPAばく露と幼児期の児のBPAばく露が関連する仮説のもとで、妊娠中の母のBPAばく露のみ、あるいは出生後の児のBPAばく露のみを測定していた場合は、「やや不適切」、両者とも測定していた場合「適切である」となる。例えば、妊婦のBPAばく露の研究であれば、妊婦という母集団において、研究対象集団のBPAばく露の変動範囲が母集団と大きくことならないと想定されるか？</p> <p>適切である：4 概ね適切である：3 やや不適切である：2 不適切である：1</p>	(1) Selection (3) Detection (exposure characterisation)
7	エンドポイント	2~4	<p>健康影響に関するエンドポイントが明確に定義されているか</p> <p>疾患：4 症状：3 検査異常値（高血圧、肥満、新生児体重など定義があるもの、定義されずに用いられているものも含む）：2（定義ありと定義なしの境界が不明瞭）</p>	(4) Detection (outcome assessment)
8	エンドポイント測定方法	1~4	<p>1. 疾病の有無を把握する者（自己申告の場合はその人）に対して、ばく露の有無の情報は開示されていない。</p> <p>2. 発症（コホート研究の場合）の把握、症例と対照の把握（症例対照研究の場合）は、確立した方法によって行われている</p> <p>3. エンドポイントの測定は、ばく露の間の時間間隔がそのエンドポイントの特徴に照らし合わせて適切である。</p> <p>1~3 全てを満たしている：4 1~3 全てを満たしていないが、そのことによる研究結果の偏りはないと想定：3 1~3 のいずれかに問題があることが想定されるか、記載が十分でない場合：2 1~3 のいずれかに明らかな問題がある：1</p>	(4) Detection (outcome assessment)

	9	交絡の制御	1~4	<p>1. 主要な交絡要因を標準化、マッチング、多変量解析、層化、傾向スコア分析などの適切な統計学的方法を用いて補正している。</p> <p>2. 交絡要因の測定方法は信頼性・妥当性の高い方法によっている。 1~2に明らかに該当し、適切である:4 1と2に部分的に該当し、概ね適切である:3 記載がない、やや不適切で残余交絡の可能性がある:2 不適切である:1</p>	(5) Confounding
	10	統計学的手法	1~4	<p>以下に示すような統計学的方法についての詳細が十分に説明されている。また統計学的方法の内容は適切であり、一般的かつよく使われるものである。</p> <ul style="list-style-type: none"> 元データを変換するための準備段階の分析の詳細が示されている（実施の場合） 主要な分析に用いられる変数が明示され、記述統計量が示されている 研究の主要な目的のための分析方法について十分解説されている 実際のデータが解析手法の前提条件となるデータの分布等に合致していることが示されている 多重比較のための適切な統計学的手法で実施されている 外れ値の取り扱いが示されている 片側か両側検定かが明記され、片側の場合はその使用の正当化理由が示されている αエラーの水準が示されている 統計手法の引用文献が示されている 使用した統計解析ソフトが示されている。 信頼区間やp値が提示されている <p>適切である:4 概ね適切である:3 やや不適切である:2 不適切である:1</p>	(7) Other sources of bias
結果	11	結果報告	1~4	<p>観察指標は欠落なく報告されているか、すなわち測定されるエンドポイント（最終エンドポイント、中間エンドポイント）は全て方法、抄録、イントロダクションに記載され、実際に報告されている。</p> <p>適切である:4 概ね適切である:3 やや不適切である:2 不適切である:1</p>	(6) Selective reporting

表 7. 各文献が報告するエンドポイントとそれをグループ化したエンドポイントのクラスターとの対応

各文献のエンドポイント	証拠の信頼度を評価するエンドポイントのクラスター
ALT	Liver toxicity
AST	Liver toxicity
gamma-glutamyl transpeptidase (γ -GTP)	Liver toxicity
Hyperuricemia	Cardiometabolic effects*
Acute bronchiolitis	Asthma/allergy
Acute otitis media	Asthma/allergy
Allergic diseases	Asthma/allergy
Asthma	Asthma/allergy
Atopic dermatitis	Asthma/allergy
Atopy	Asthma/allergy
Bacterial colonisation	Asthma/allergy
Bronchitis	Asthma/allergy
Chest infections	Asthma/allergy
Croup	Asthma/allergy
Eczema	Asthma/allergy
Forced respiratory volume in 1s (FEV1)	Asthma/allergy
Fraction of exhaled nitric oxide (FENO)	Asthma/allergy
IgE	Asthma/allergy
PC20	Asthma/allergy
Pneumonia	Asthma/allergy
'Rashes, eczema or hives'	Asthma/allergy
Rhinitis (allergic)	Asthma/allergy
Wheeze	Asthma/allergy
Infectious enteritis	Infections other than respiratory tract
Urinary tract infections	Infections other than respiratory tract
Adiponectin	Obesity
Birth weight (BW)	Obesity
Body Fat	Obesity
Body mass index (BMI)	Obesity
Fat leg	Obesity
Fat mass	Obesity
Fat mass index (FMI)	Obesity

Fat trunk	Obesity
Fat trunk/leg ratio	Obesity
Leptin	Obesity
Obesity	Obesity
Obesity (central)	Obesity
Overweight	Obesity
Rapid growth	Obesity
Skinfold thickness	Obesity
Subcutaneous adipose tissue (SAT)	Obesity
Visceral adipose tissue (VAT)	Obesity
Visceral/Subcutaneous adipose tissue ratio (VAT/SAT)	Obesity
Waist circumference	Obesity
Weight change	Obesity
Weight gain (annual)	Obesity
Cholesterol (total)	Cardiometabolic effects
High density lipoprotein (HDL)	Cardiometabolic effects
Low-density lipoprotein (LDL)	Cardiometabolic effects
Triglycerides	Cardiometabolic effects
Diastolic blood pressure	Cardiometabolic effects
Pulse Pressure	Cardiometabolic effects
Systolic blood pressure	Cardiometabolic effects
Free thyroxine (FT4)	Thyroid
Free triiodothyronine (FT3)	Thyroid
Thyroid stimulating hormone (TSH)	Thyroid
Total thyroxine (TT4)	Thyroid
Total triiodothyronine (TT3)	Thyroid
C-peptide	Type 2 diabetes mellitus (T2DM)
C-peptide index (CPI)	Type 2 diabetes mellitus (T2DM)
Type 2 diabetes mellitus (T2DM)	Type 2 diabetes mellitus (T2DM)
Epidermal growth factor receptor (eGFR)	Type 2 diabetes mellitus (T2DM)
Glucose	Type 2 diabetes mellitus (T2DM)
Homeostasis model assessment of insulin resistance (HOMA-IR)	Type 2 diabetes mellitus (T2DM)
IGF-1	Type 2 diabetes mellitus (T2DM)
Insulin	Type 2 diabetes mellitus (T2DM)

Leptin	Type 2 diabetes mellitus (T2DM)
Gestational diabetes mellitus (GDM)	Gestational diabetes mellitus
Glucose	Gestational diabetes mellitus
Impaired glucose tolerance	Gestational diabetes mellitus
Anxiety	Neurodevelopment
Behaviour	Neurodevelopment
Behaviour (Behaviour Assessment System for Children-2, BASC-2)	Neurodevelopment
Behaviour (boys)	Neurodevelopment
Behaviour (Behaviour Rating Inventory of Executive Function-Preschool, BRIEF-P)	Neurodevelopment
Behaviour (girls)	Neurodevelopment
Cognitive development	Neurodevelopment
Cortisol	Neurodevelopment
Cortisol reactivity	Neurodevelopment
Depression	Neurodevelopment
Executive function	Neurodevelopment
Intellectual ability	Neurodevelopment
Psychomotor development	Neurodevelopment
Social responsiveness scale-2 (SRS-2)	Neurodevelopment
Sociality	Neurodevelopment
Sociality - Autistic behaviour (Social Responsiveness Scale)	Neurodevelopment
Virtual Morris water maze (VMWM)	Neurodevelopment
Abdominal circumference	Fetal and post-natal growth
Biparietal diameter	Fetal and post-natal growth
Birth length	Fetal and post-natal growth
Birth weight	Fetal and post-natal growth
Chest circumference	Fetal and post-natal growth
Femoral length	Fetal and post-natal growth
Gestational age at birth	Fetal and post-natal growth
Growth rate	Fetal and post-natal growth
Head circumference	Fetal and post-natal growth
Height	Fetal and post-natal growth
Large for gestational age	Fetal and post-natal growth
Low birth weight (LBW)	Fetal and post-natal growth

Length of gestation	Fetal and post-natal growth
Placental weight	Fetal and post-natal growth
Ponderal index	Fetal and post-natal growth
Small for gestational age (SGA)	Fetal and post-natal growth
Weight	Fetal and post-natal growth
Weight for length	Fetal and post-natal growth
Weight for length (birth)	Fetal and post-natal growth
Weight (fetal)	Fetal and post-natal growth
Gestational age at birth	Prematurity
Length of gestation	Prematurity
Preterm delivery	Prematurity
Pre-eclampsia (onset, severity)	Pre-eclampsia
E2	Male fertility
Estrone	Male fertility
Fecundability	Male fertility
Follicle stimulating hormone (FSH)	Male fertility
FSH:Inhibin B ratio	Male fertility
Hypo-osmotic swollen	Male fertility
Inhibin B	Male fertility
LH:Testosterone ratio	Male fertility
Live Birth rate	Male fertility
Luteinising Hormone (LH)	Male fertility
Sex	Male fertility
Sperm amplitude head displacement	Male fertility
Sperm average path velocity	Male fertility
Sperm beat cross frequency	Male fertility
Sperm concentration	Male fertility
Sperm count	Male fertility
Sperm curvilinear velocity	Male fertility
Sperm DNA fragmentation index	Male fertility
Sperm head acrosome area	Male fertility
Sperm head area	Male fertility
Sperm head elongation factor	Male fertility
Sperm head length	Male fertility
Sperm head perimeter	Male fertility
Sperm head width	Male fertility

Sperm high DNA stainability	Male fertility
Sperm linearity	Male fertility
Sperm morphology amorphous	Male fertility
Sperm morphology bicephalic	Male fertility
Sperm morphology coiled tail	Male fertility
Sperm morphology cytoplasmatic droplet	Male fertility
Sperm morphology megalo head	Male fertility
Sperm morphology micro head	Male fertility
Sperm morphology neck and midpiece abnormalities	Male fertility
Sperm morphology other tail abnormalities	Male fertility
Sperm morphology pyriform	Male fertility
Sperm morphology round	Male fertility
Sperm morphology strict criteria	Male fertility
Sperm morphology taper	Male fertility
Sperm morphology traditional normal	Male fertility
Sperm motility	Male fertility
Sperm straight line velocity	Male fertility
Sperm straightness	Male fertility
Sperm straw distance [=motility]	Male fertility
Sperm volume	Male fertility
Testicular volume	Male fertility
Testosterone	Male fertility
Aneuploid pregnancy	Female fertility
Antral follicle count	Female fertility
Clinical pregnancy	Female fertility
Corpus luteum rescue	Female fertility
Day-3 FSH	Female fertility
Early miscarriage	Female fertility
Embryo quality	Female fertility
Endometrial wall thickness	Female fertility
E2	Female fertility
E2 trigger level	Female fertility
Fecundability	Female fertility
Fertilisation rate	Female fertility
FSH	Female fertility
Follicular phase length	Female fertility

Human chorionic gonadotropin	Female fertility
Implantation	Female fertility
Infertility	Female fertility
Live birth rate	Female fertility
Luteal phase length	Female fertility
Luteinising hormone (LH)	Female fertility
Metaphase II (MII) oocytes	Female fertility
Miscarriage	Female fertility
Ovarian volume (OV)	Female fertility
probability of implantation	Female fertility
Progesterone	Female fertility
Sex	Female fertility
Time to implantation	Female fertility
Total oocytes	Female fertility
% Mammary fibroglandular volume (%FGV)	Pubertal/endocrine
2D:4D digit ratio	Pubertal/endocrine
Androstenedione	Pubertal/endocrine
Anogenital distance anus - penis (AGDap)	Pubertal/endocrine
Anogenital distance anus - clitoris (AGDac)	Pubertal/endocrine
Anogenital distance anus - fourchette (AGDaf)	Pubertal/endocrine
Anogenital distance anus - scrotum (AGDas)	Pubertal/endocrine
Breast Development	Pubertal/endocrine
Breast volume	Pubertal/endocrine
Breastfeeding	Pubertal/endocrine
Dehydroepiandrosterone (DHEA)	Pubertal/endocrine
Dehydroepiandrosterone Sulfate (DHEA-S)	Pubertal/endocrine
E2	Pubertal/endocrine
Facial hair growth	Pubertal/endocrine
Free testosterone	Pubertal/endocrine
Genital development	Pubertal/endocrine
Gonadarche	Pubertal/endocrine
Inhibin B	Pubertal/endocrine
Luteinising hormone (LH)	Pubertal/endocrine
Mammary fibroglandular volume (FGV)	Pubertal/endocrine
Menarche onset	Pubertal/endocrine
Perceived insufficient milk supply (PIM)	Pubertal/endocrine

Peripubertal hormone levels	Pubertal/endocrine
Pubarche (boys)	Pubertal/endocrine
Pubarche (girls)	Pubertal/endocrine
Pubertal development scale	Pubertal/endocrine
Pubic Hair	Pubertal/endocrine
Sex hormone-binding globulin (SHBG)	Pubertal/endocrine
Testicular volume	Pubertal/endocrine
Testosterone	Pubertal/endocrine
Thelarche	Pubertal/endocrine
Voice change	Pubertal/endocrine
Prostate cancer	Prostate
Lymphoma	Lymphoid tissues

* Hyperuricemia は、EFSA では Kidney toxicity にグループ化されたが、ここでは Cardiometabolic effect に分類した。

表 8. エンドポイントのクラスターごとの証拠の信頼度の評価対象論文数

Health Outcome Categories (HOC)	Endpoint cluster	論文数 (コホート、症例対照に限定)
General toxicity	Liver toxicity	1
	Kidney toxicity	3
	Cardiotoxicity	1
Effects involving the immune system	Asthma/allergy	12
	Infections other than respiratory tract	1
Metabolic and related endocrine effects	Obesity	6
	Cardiometabolic effects	4
	Thyroid	3
	Type 2 diabetes mellitus (T2DM)	9
	Gestational diabetes mellitus	5
Mental and neurological effects (Neurotoxicity)	Neurodevelopment	23
Reproductive and developmental toxicity	Fetal and post-natal growth	8
	Prematurity	8
	Pre-eclampsia	3
	Male fertility	8
	Female fertility	10
	Pubertal/endocrine	40
Carcinogenicity	Prostate	2
	Others	4
Total		151

表 9. エンドポイントのクラスター別の証拠の信頼度（男性妊孕性）

	Initial Rating 初期評価	Downgrading				Upgrading				エビデンスの総合的評価
		Risk of Bias 文献の信頼性評価が低い	Unexplained inconsistency 研究間の理由のつかない結果の不一致	Indirectness 代替指標、代替的对象集団での検討による間接的な知見 外的妥当性がない	Imprecision 推定された関連性指標の精度が低い	Magnitude 強い関連性	Dose response 量反応関係がある	Residual confounding 残余交絡	Consistency 異なる研究デザイン・研究対象集団での結果の一致	
Male fertility 男性の妊孕性										
Male fertility 男性の妊孕性文献の 細分類	高い 中等度 低い	⊖⊖ ⊖ ±0	⊖⊖ ⊖ ±0	⊖⊖ ⊖ ±0	⊖⊖ ⊖ ±0	±0 ⊕ ⊖⊖	±0 ⊕ ⊖⊖	±0 ⊕ ⊖⊖	±0 ⊕ ⊖⊖	高い 中等度 低い
男性不妊 (6 文献) コホート：323, 353, 371, 372 症例対照：175, 218	中等度	±0 全て問題なし	±0 323, 353, 371, 372, 175, 218 で関連なし	±0 不妊治療を受けている男性 全て問題なし	±0 効果の上限 ／下限比、 差	±0 323, 353, 371, 372, 175, 218 で関連性なし	±0	±0	⊖⊖ コホートと 症例対照研究の両者で 一致した結果	中等度または高い (関連なし)
精液の質 (4 文献) コホート：337, 1285 症例対照：175, 218	低い 337 は高い	±0 全て問題なし	±0	±0 不妊治療を受けている男性 全て問題なし	±0 効果の上限 ／下限比、 差	±0 175, 218 で関連性なし	⊕ 1285 (連続量での 解析)で量 反応関係の可能性	±0	±0	低い (関連あり)

※175、218、323、337、353、371、372、1285 は文献番号

表 10. エンドポイントのクラスター別の証拠の信頼度（女性妊孕性）

Female fertility 女性の妊孕性	Initial Rating	Downgrading				Upgrading				エビデンス の総合的評 価
		Risk of Bias	Unexplained inconsistency	Indirectness	Imprecision	Magnitude	Dose response	Residual confounding	Consistency	
Female fertility 文献 の細分類	高い 中等度 低い	⊖ ⊖ ±0	⊖ ⊖ ±0	⊖ ⊖ ±0	⊖ ⊖ ±0	±0 ⊕ ⊕⊕	±0 ⊕ ⊕⊕	±0 ⊕ ⊕⊕	±0 ⊕ ⊕⊕	高い 中等度 低い
体外受精 (IVF) の結果 (5 文献)、妊 娠までの期間 (1 文献) コホート: 44, 253, 316, 377 (後三者は同じ EARTH Study)、412, 1304	中等度	±0 全て問題 なし	⊖ 377(栄養素で 層化しない全 体)で関連なし だが、253, 316のように 栄養素摂取に よる効果修飾 あり、その理 由は必ずしも 明確でない。	⊖ 不妊治療中の 女性における 結果 253 は葉酸、 316 は大豆低 摂取群での関 連で外的妥当 性は低い	±0 全て imprecision の基準は満 たさない	±0 44 でのみ 強い関連	±0 44、412 は連 続量で解析だ が量反応関係 は不明、253, 3167 は 4 分 位での解析で 第 4 四分位 vs. 第 1~3 分 位の間でのみ 関連	±0	±0	低い (効果修飾 の結果は解 釈が難しい が全体とし て関連なし)
卵巣機能(卵母 細胞の成熟) (4 文献) コホート: 167, 399 症例対照: 210, 1278	低い	±0 全て問題 なし 210 では 交絡調整 なし	±0	±0 不妊治療中の 女性における 結果	±0 全て imprecision の基準は満 たさない	±0 1278 で関 連性なし	±0 167 (連続量 での解析)で 量反応関係の 可能性、210, 399 はなし	±0 210 は交絡 調整なし	⊕ コホートと 症例対照研 究で一致し た結果	低い (関連あり)

※ 44、167、210、253、316、377、399、412、1278、1304 は文献番号