

海外のリスク評価機関における評価結果等に関する調査ー

3. (4) 残留農薬のリスク評価にあたっての毒性試験結果の解釈に係る知見の収集

① 各種毒性試験における毒性を解釈する上で共通する考え方 (適応性変化等)

番号	著者名	文献名	雑誌：巻：ページ	参照箇所の該当ページ
1	World Health Organization (WHO)	Guidance Document for WHO Monographers and Reviewers	2015	Part3: 1.1 (34), 1.2 (34), 1.3 (34-36), 1.4 (36), 1.10 (42)
2	Lewis RW, Billington R, Debryune E, Gamer A, Lang B, Carpanini F	Recognition of Adverse and Nonadverse Effects in Toxicity Studies	Toxicol Pathol. 2002; 30(1): 66-74	Abstract, Figure2
3	Williams GM, Iatropoulos MJ	Alteration of Liver Cell Function and Proliferation: Differentiation Between Adaptation and Toxicity	Toxicol Pathol. 2002; 30(1): 41-53	Abstract
4	Perry R, Farris G, Bienvenu JG, Dean C Jr, Foley G, Mahrt C, Short B	Society of Toxicologic Pathology Position Paper on Best Practices on Recovery Studies: The Role of The Anatomic Pathologist	Toxicol Pathol. 2013; 41(8): 1159-1169	Abstract, 1164
5	Hall AP, Elcombe CR, Foster JR, Harada T, Kaufmann W, Knippel A, Küttler K, et al	Liver Hypertrophy: A Review of Adaptive (Adverse and Non-adverse) Changes--Conclusions from the 3rd International ESTP Expert Workshop	Toxicol Pathol. 2012; 40(7):971-994	Abstract

番号	1-01
著者名	World Health Organization (WHO)
文献名	Guidance Document for WHO Monographers and Reviewers
巻, ページ, 発行年	2015
分類 (①~④)	①
参照箇所の該当ページ	Part3: 1.1 (34), 1.2 (34), 1.3 (34-36), 1.4 (36), 1.10 (42)

概要

ヒトの健康に対する化学物質の有害影響を判断する際に重要となる毒性試験から得られる量的成果に対して、統計学的あるいはそれ以外の観点も含めて意義のある有害影響を起こさない最高用量/濃度を決定することが重要である。そのような毒性試験からの所見に対する首尾一貫した解釈が重要であり、それはこれらのデータがヒトの健康リスク評価、特に「安全な」ばく露レベルを決定する重要な意味を有しているためである。したがって、毒性試験において観察された影響が毒性か否かを判断することはヒトの健康リスク評価において最も重要かつ、論争上の課題の一つである。観察された影響が一過性、可逆性あるいは軽度な場合、恒常性を損なわない限り、毒性ではないと判断する。

肝臓の肥大の毒性学的解釈については、病理組織学的な障害性変化（壊死、炎症、線維化等）または生化学的に肝障害を示す変化（AST, ALT, ALP, ビリルビンの高値等）を伴わない場合は適応性変化と考える。肝臓の肥大が適応反応か毒性かを判断する情報が不十分な場合、あるいは追加的な負荷に対応する機能障害をきたしうる場合（例：累積的影響）や他の化学物質の影響に対する感受性の増加をもたらす可能性がある場合は、その影響は有害影響と仮定することをデフォルト、すなわち標準的な方針とする。設定された参照値（例：ADI）は薬物代謝酵素誘導の LOEL を超えてはならない。

統計学的に有意差があっても生物学的に意義がない影響がある。一方で、軽微ではあるが生物学的に意義のある変化あるいは低用量での毒性影響は検出力の低い試験法では検出できない可能性がある。“統計的に有意”とされる変化は、その影響の重要性または生物学的関連性と必ずしも直接結びつくわけではない。軽微、軽度、中等度、重度に分類された組織学的所見などの実験データを考察する場合、程度を考慮した統計解析を行うことで、総発生数だけでなく、程度についても用量反応性の有無を調べることが出来る。統計学的に有意な所見は、投与との関連を示唆するが、それをそのまま生物学的な毒性影響とか、毒性学的に意義ある影響とは結論付けられない。実際、正常な生物学的変動の範囲内であっても統計学的には有意差があると判定される可能性がある（例：再生反応の兆候がない赤血球数の減少）。これに対して、生物学的あるいは毒性学的に検討することで統計学的な結果が覆ることがある（例：1群4匹のみで実施するイヌ試験の場合）。しかし、統計学的に有意でない結果に対して生物学的意義を主張する場合には常に注意が必要である。統計学的な有意差はないが用量との関連性のある軽度な病理組織所見は、投与と関連があるか否か、毒性か否かを判断するのが難しい場合がある（例：肝細胞内の空胞の軽度な増加、脾臓における髄外造血の軽度な増加は毒性学的に重要とはみなさなくてよい）。全ての毒性試験において LOAEL で認められる軽度な所見は、ケースバイケースで包括的な議論が必要になる。

投与期間中に消失する影響は一過性の影響として一般的に認められ、次のようなことに起因しうる。

1) 投与に対する非特異的反応（例：投与開始直後の摂餌忌避）、2) 強制経口投与によるストレスが原

因の体重減少，3) 年齢や体重増加に伴う摂餌量の減少。

投与期間中に消失する一過性の影響と投与終了後に回復する影響とを区別する必要がある。前者については，急性ばく露でみられ，投与期間中に消失した一過性の影響が有害影響である可能性があるため急性参照用量（ARfD）の設定に重要であり，後者の投与期間終了後に回復する影響については，少なくとも継続的ばく露の場合では，投与期間中に現れる影響が有害影響である可能性があるため，慢性参照用量を設定するためにそれぞれ重要である。

体重に対する一過性の影響（例：餌の嗜好性の変化による摂餌量の低下，体重減少が起こるが実験動物の適応とともに影響が消失する場合）は毒性とは考えない。しかし，この結論は動物の成長の抑制が餌の嗜好性低下によって直接もたらされた影響であることを証明することが可能な場合にのみ導き出されるということが重要である。

番号	1-02
著者名	Lewis RW, et al
文献名	Recognition of Adverse and Nonadverse Effects in Toxicity Studies
巻, ページ, 発行年	Toxicol Pathol.2002; 30(1): 66-74
分類 (①～④)	①
参照箇所の該当ページ	Abstract, Figure2

毒性試験から得られる最も重要な定量的アウトプットの一つは、ヒトの健康リスク評価に妥当である (relevant) と考えられる投与に関連した影響を引き起こさない最も高いばく露レベル (用量または濃度) の特定である。しかし、規制当局及び他の科学文献ならびに現状の実務をレビューした結果、下記のような頻繁に使用される用語の定義や適用に一貫性がないことが判明した。さらに、有害影響と有害でない影響の評価や鑑別を含め、毒性試験結果の解釈の指針となるような首尾一貫した基準は認められなかった。

そこで、まずは毒性試験の結果を記述するために頻繁に使用される主要用語の定義を提案する。

無作用量 (NOEL): 適切な対照群と比較して、ばく露群において (有害・有害でない) 影響が観察されない最高のばく露量。

無毒性量 (NOAEL): 適切な対照群と比較して、ばく露群において有害影響の頻度・重症度に統計学的あるいは生物学的に有意な増加が認められない最高のばく露量。このばく露量において何らかの影響が生じる可能性はあるが、それらは有害影響や有害影響の前駆状態 (precursor) とは考えない。

最小毒性量 (LOAEL): 適切な対照群と比較して、ばく露群において有害影響の頻度・重症度に統計学的あるいは生物学的に有意な増加が認められる最低のばく露量。

有害影響: 単独あるいは複合的に生体全体のパフォーマンスに有害影響を及ぼす、あるいは環境変化への生体能力を低下させる、(刺激に対する) 生化学的、形態学的あるいは生理学的変化。反対に、有害でない影響は、生体の全身状態、成長、発達、寿命に影響を与えるような生化学的、形態学的あるいは生理学的変化を引き起こさない生物学的影響と定義できる。

生物学的・毒性学的に意義のある影響: 生体システムの健康に (正または負の) 重要あるいは顕著な影響を与えると考えられる (刺激に対する) 反応。この概念は、(生体の全身状態に対しては意義のある場合とない場合がある) 統計学的に有意な影響・変化とは区別されるべきである。

次に、毒性学者が一貫性のある試験結果の解釈を行なうことを支援する首尾一貫したフレームワーク (図 1 参照) を概説する。このフレームワークのプロセスには以下の 2 つの主要なステップがあり、それぞれにおける一貫した判断を行なうための基準が記載されている。

1) 対照群の値との差が投与の影響によるものか、偶発性のものかを判断する。

次の場合は投与による影響の可能性は低い: 明らかな用量反応関係性がみられない。差の原因が外れ値にある。測定精度が悪い。正常な生物学的変動の範囲内である。生物学的妥当性に欠ける。

2) 投与による影響であると判断された変化についてはさらに評価し、有害影響であるかどうかを判断する。

次の場合は有害影響の可能性は低い: 生体あるいは影響を受けた臓器・組織の機能障害がない。適応反応である。一過性のものである。重篤度が低い。影響が単独発現している (その影響に通常関

連する他のパラメータの変化を伴わない)。有害影響に進行する前駆状態ではない。他の有害影響の二次的影響である。影響が投与方法に起因する。

複雑な試験結果を解釈するには、これらの基準を組み合わせることで総合的な判断を下す証拠の重み付けアプローチが最適な方法であると考えられる。そして、このようなスキームを使用することで、ハザードとリスク評価の基礎となる試験結果の解釈について一貫性を向上させることができると考えられる。

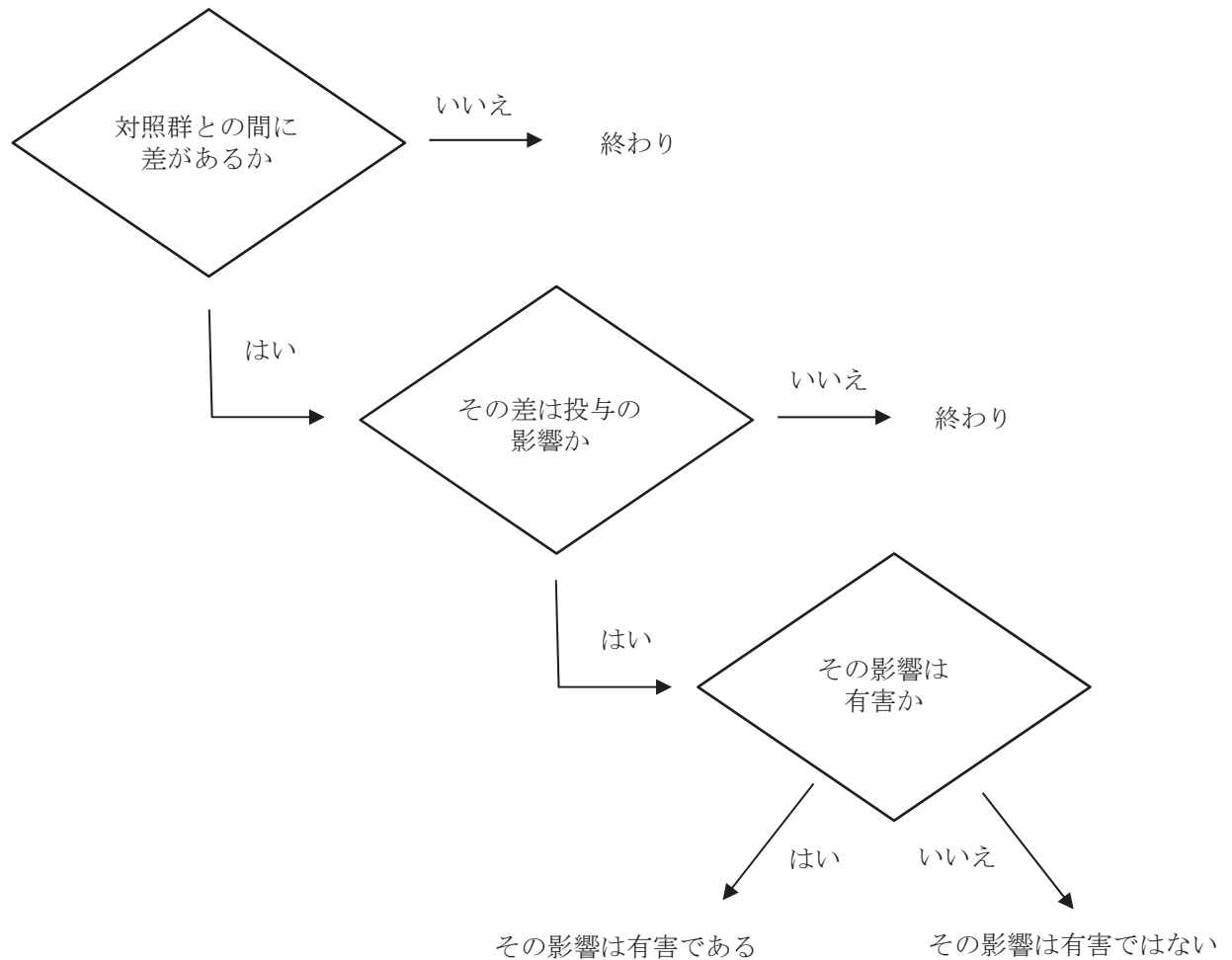


図1—毒性試験結果の解釈のためのフレームワーク

番号	1-03
著者名	Williams GM, Iatropoulos MJ
文献名	Alteration of Liver Cell Function and Proliferation: Differentiation Between Adaptation and Toxicity
巻, ページ, 発行年	Toxicol Pathol. 2002; 30(1): 41-53
分類 (①~④)	①
参照箇所の該当ページ	Abstract

実験動物が生物学的な影響の出るレベルの化学物質 [内因性あるいは外因性 (合成あるいは天然由来)] にばく露されると、様々な組織単位で化学的摂動 (変化またはかく乱) に対処するための反応が誘発される。ある種の化学物質では、初期反応は適応反応を構成し恒常性を維持するが、どの組織レベルにおいてもそれが崩れると毒性 (有害) 影響が生じる。

実験動物やヒトの肝臓は、他の臓器と同様に、細胞の増殖に続く分化を特徴とするプログラムされた段階的な成長、発達を経るが、分化した細胞は恒常性を維持するための機能及び生体のニーズに応えるための特殊化した機能を持つようになる。成熟後の肝臓は、肝細胞 (中間細胞に分類され、複製頻度は低いもののシグナルに応答して複製することができる)、複製細胞である幹細胞、内皮細胞、クッパー細胞、星細胞 (伊東細胞, 周皮細胞)、胆管上皮及び顆粒リンパ球 (ピット細胞) で構成される。これらの様々な細胞において、化学物質による分子レベルの定量可能な変化が細胞小器官レベルでの変化の根底にあり、細胞小器官レベルでの変化は細胞レベルでの変化を誘導し、最終的には全身性の変化につながる。この定量可能な変化には、細胞レベルで生じる複製、壊死・アポトーシス、分化のように計数的 (2 値的, すなわち **all or none**) なものがある一方、細胞内あるいは細胞外レベルで生じる酵素誘導、細胞小器官の肥大、細胞外基質の生成のように、段階的あるいは連続的 (非 2 値的 **nonbinary**) なものもある。そして、あらゆるこのような変化によって反応や影響は構成される。肝臓におけるそれぞれの種類の細胞は、その局在と機能に応じて刺激に反応するが、一般的に再生を行なっている細胞は、ほぼ静止している中間細胞と比較して化学的損傷に対しより脆弱である。

肝臓の適応反応では通常、化学物質が (多くの場合受容体を介して) 細胞の制御経路に作用することで、遺伝子発現が変化し、最終的にメタボローム (**metabolome**) が変化する。この反応は、様々な細胞及び細胞外機能の調節を通じて、恒常性を維持しようとするためのものである。また適応反応は、あらゆる組織レベルで化学物質によるストレスへの応答能力を高め、可逆的であり、生存能力を維持するという点で有益である。毒性影響が生じる閾値を下回るばく露でのこのような適応反応は、ホルミシス (**hormesis**) とも呼ばれている。一方、肝臓における毒性 (有害) 影響では、化学物質が細胞の高分子と反応することで、恒常性の崩壊が起こる。このような影響は、ストレスへの応答能力を低下させ、非可逆的となり得、また生存能力を低下させ得る。通常は適応反応を誘導するレベルのばく露量でも、ばく露期間の長期化、もしくは高用量のばく露により閾値を上回った場合は、用量の変化につれて作用機序が変化し毒性影響が現れうる。従って、影響を及ぼす用量の変化とともに作用機序も変化していることになる。肝臓では様々な適応反応及び毒性影響が確認されており、前者を誘導するものとしてフェノバルビタールやシプロフィブラートが、後者を誘導するものとして *p*-ジクロロベンゼンや 2-アセチルアミノフルオレンが挙げられる。肝臓への化学物質の影響は、例外はあるものの概して実験

動物とヒトで類似しているため、適応反応と毒性影響の鑑別とモニタリングは安全性評価上不可欠である。

番号	1-04
著者名	Perry R, et al
文献名	Society of Toxicologic Pathology Position Paper on Best Practices on Recovery Studies: The Role of the Anatomic Pathologist
巻, ページ, 発行年	Toxicol Pathol. 2013; 41(8): 1159-1169
分類 (①～④)	①
参照箇所の該当ページ	Abstract, 1164

病理学者は、毒性試験において認められた病理組織学的所見の回復性を予測するためには、組織の種類、細胞外基質の関与及び病変形成過程の性質ならびに重症度を考慮すべきである。

組織は、その構成細胞の増殖能に応じて不安定組織、安定組織及び永久組織のいずれかに分類される。不安定組織は高い回復力を有し、細胞外基質（細胞を囲む間質基質と基底膜で構成されるネットワークであり、水分とミネラルを隔離する機能を持ち、細胞が付着する足場を提供し、さらに増殖因子の貯蔵庫としての役割を果たす）が無傷である限り、同種の細胞の複製及び再増殖による再生を行なう。安定組織は、不安定組織と同様に反応及び再生ができるものの、その程度は低い。また、正常な機能や修復のためには、細胞外基質が正常・無傷であることが必要である。細胞外基質の間質フレームワークが損傷した場合、再生した実質細胞が臓器内に不規則に散乱し、効果的な修復ができず、再生がうまくいかない。永久組織が致命的な傷害を受けた場合、同種の細胞による置換は行なわれず、通常は結合組織で修復される。これらの組織の増殖（再生）能とその例を以下にまとめた。

1. 不安定組織（絶えず細胞分裂している組織）
 - a. 細胞は生涯にわたって増殖する
 - b. 傷害後、容易に再生する
 - c. 幹細胞を含んでいる
 - d. 例：骨髄、造血組織、表面上皮（皮膚、消化管上皮）
2. 安定組織（静止状態の組織）
 - a. 低レベルの複製
 - b. 元の組織を再生・再構築することができる
 - c. 刺激に反応して急速に分裂することができる
 - d. 例：肝臓、腎臓、膵臓の実質細胞；線維芽細胞、平滑筋、内皮、リンパ球、骨細胞、筋衛星細胞、軟骨細胞などの間葉系細胞
3. 永久組織（細胞分裂しない組織）
 - a. 出生後、通常は細胞分裂しない
 - b. 再生できない
 - c. 例：神経細胞、心筋

再生過程における細胞外基質と細胞の関連性や相互作用は極めて複雑であり、例外もあるものの、

回復性の予測においては一般的に以下のことが適用できる。

1. 細胞外基質が影響を受けず，組織傷害が不安定性組織あるいは安定性組織に生じた場合，回復の見込みがある。
2. 組織傷害によって細胞外基質が損傷あるいは損失した場合は，その損傷や損失は回復せず，無秩序で機能を喪失する再生，線維化あるいは鉍質沈着が生じる可能性が高い。

番号	1-05
著者名	Hall AP, et al
文献名	Liver Hypertrophy: A Review of Adaptive (Adverse and Non-adverse) Changes-- Conclusions from the 3rd International ESTP Expert Workshop
巻, ページ, 発行年	Toxicol Pathol. 2012; 40(7): 971-94
分類 (①~④)	①
参照箇所の該当ページ	Abstract

概要

げっ歯類を用いた亜慢性（亜急性）及び慢性の毒性試験において薬物や化学物質によって引き起こされる肝臓の腫大は、同程度の用量での生涯試験における肝毒性や発がん性、あるいはヒトとの関連性についての認識の点で長年にわたり毒性学の専門家を悩ませてきた。

肝肥大とは、病理組織学者にとっては明確に認識され組織学的に十分定義されている一方、毒性学者にとっては器官重量の増加、肝細胞の平均サイズの増加（肝細胞の肥大）、そして（時に機能的な「亢進性」肥大とも言われる）肝酵素の誘導さえも含まれる意味の言葉として捉えられている。前臨床毒性試験において、実験動物に長期間（終身）肝酵素誘導剤をばく露した結果、肝重量の増加、肝細胞の肥大、細胞増殖、そして長期試験における高頻度の肝発がんといった特徴を示す毒性学的変化が生じることが報告されている。近年の進歩により、多くの外因性化学物質は、CAR, PXR, PPARα といった核内ホルモン受容体の活性化を介した共通の作用機序により、これらの変化が誘導される可能性があることが明らかになってきた。これらについて改変された核内ホルモン受容体を発現する遺伝子組み換えマウスを作製し、他の研究とも複合的に検討したところ、これらの作用の多くがげっ歯類に特異的であることが明らかになった。これらのデータは、ヒトについて得られた信頼できる利用可能な疫学的及び経験的なエビデンスと一致しており、これらの変化はヒトではほとんど起こらないという科学的見解を裏付けるものである。そのため ESTP は、肝細胞肥大に関連して何が有害と考えられるかを毒性病理学者がより明確に定義するために、国際的な専門家パネルを招集し、エビデンスについて議論した。このワークショップの結果、肝毒性を示す組織学的または臨床病理学的変化を伴わない肝細胞肥大の結果としての肝肥大は適応性であり、有害反応ではないと結論づけられた。

肝重量の増加によってもたらされる組織学的変化を評価する際、その変化が有害か否かを判断するために多くの段階を経て検討する必要がある。まず肝細胞の壊死、線維化、炎症や脂肪性の空胞変性、胆管/卵形細胞の増殖、変性、線維化、胆汁うっ滞、他の肝臓構成細胞の壊死や変性といった構造的変性や壊死性変化の組織学的証拠があるかを検討する。組織学的変化がない場合、ALT の少なくとも 2 倍から 3 倍の増加や他の肝胆管傷害の指標（ALP, AST, γ GT, GLDH 等）の生物学的に有意な変動。さらに別の肝臓の異常を示す臨床検査マーカー（アルブミン、ビリルビン、胆汁酸、凝固因子、コレステロール、過酸化脂質等）の生物学的に有意な変動のうち、少なくとも 2 つの肝臓の検査項目が用量依存的であり生物学的に有意かつ継続的に変動しているとして特徴付けられる肝細胞傷害の臨床病理学的証拠があるかを検討する。以上述べた有害作用の基準に達する変化が認められない場合、臓器重量の増加や肝細胞肥大が薬物代謝酵素誘導に関連した適応反応であり、ヒトには毒性学的意義がないと判断しうる段階となる。その発現メカニズムに AhR*や既知の毒性メカニズムが関与している場合は

有害作用と判断され、CAR, PXR, PPARa の活性化**あるいは未知のメカニズムであってもその他の毒性メカニズムを示す証拠がない場合、有害作用とはみなされないと判断する。

*げっ歯動物に遺伝子毒性がないダイオキシンを与えると、AhR の活性化を通じた肝肥大や肝細胞の変異増殖巣、及び肝発がんが認められる。AhR 介在性の発がんのヒトへの外挿性は否定できない。

**CAR, PXR, PPARa の活性化はげっ歯動物の肝細胞を増殖に向かわせる。ヒトでこれら核内受容体の活性化が起こったとしてもヒトの肝細胞の増殖性は惹起され難い。

海外のリスク評価機関における評価結果等に関する調査ー

3. (4) 残留農薬のリスク評価にあたっての毒性試験結果の解釈に係る知見の収集

② 毒性判断に用いる統計学的手法、背景データの利用

番号	著者名	文献名	雑誌：巻：ページ	参照箇所の該当ページ
1	World Health Organization (WHO)	Guidance Document for WHO Monographers and Reviewers	2015	Part 2: 6.2.6 (16), Part 3: 1.5 (36,37), 1.6 (37,38)
7	OECD	Guidance Document 116 on The Conduct and Design of Chronic Toxicity and Carcinogenicity Studies, Supporting Test Guidelines 451,452 and 453--2nd Edition. ENV/JM/MONO(2011)47	2012	4.4 (115,116)
8	Mann PC, Vahle J, Keenan CM, Baker JF, Bradley AE, Goodman DG, Harada T, Herbert R, et al	International Harmonization of Toxicologic Pathology Nomenclature: An Overview and Review of Basic Principles	Toxicol Pathol. 2012; 40(4 Suppl):7S-13S	9S-11S, Table 1
9	Elmore SA, Peddada SD	Points to Consider on the Statistical Analysis of Rodent Cancer Bioassay Data When Incorporating Historical Control Data	Toxicol Pathol. 2009; 37(5): 672-676	Abstract
19	OECD	Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies	2002	27,28,36, Appendix V

番号	2-01
著者名	World Health Organization (WHO)
文献名	Guidance Document for WHO Monographers and Reviewers
巻, ページ, 発行年	2015
分類 (①～④)	②
参照箇所の該当ページ	Part 2: 6.2.6 (16), Part 3: 1.5 (36, 37), 1.6 (37, 38)

概要

背景データ (Historical control data) は、腫瘍や発生異常の評価に利用することが多いが、それらの結果が正常な生物学的範囲内に収まることを示すことが目的である。背景データは、比較対象とする試験と同じ系統の動物を対象とし、ほぼ同時期に、同じ試験施設で実施した試験から得られたものでなければならない (理想的には、試験開始2年前から試験終了後2年までの間)。これらの基準に合わない場合は、その旨を本文中に記載する。可能なら、各試験の結果を別々に見ることが出来るような形式で提供する必要がある。最低限、試験数、平均値、最小値～最大値の範囲を記載する。当該試験と背景データの間で、解釈上または検査上の方法に違いがないことも確認する必要がある。提出された背景データがこれらの基準を満たさないなら、その不足部分を補うよう、対応を依頼する必要がある。

長期試験では、当該試験の対照動物における非腫瘍性あるいは腫瘍性病変の発生率が正常より低い場合、投与群におけるそれら病変の発生率が有意な高値を示すことがある。逆に、正常より高い場合は投与群における毒性あるいは発がん性を隠してしまう。背景データは、同じ試験施設の同じ条件下で飼育された同じ種及び系統の動物を使って行われた試験から得られたものでなければならない。背景データには次の情報が示される必要がある。

- 1) 動物種、系統の同定、生産者名、繁殖地の識別 (生産者が複数の地理的場所にある場合)
- 2) 試験施設の名称及び試験実施日
- 3) 動物の一般飼育条件 (飼料の種類及び銘柄、可能なら消費量)
- 4) 試験開始時及び計画解剖時または死亡時の対照動物の大体の年齢 (日数) 及び体重
- 5) 試験中または試験終了時に観察された対照群動物の死亡パターン及び他の関連する所見 (例: 病気、感染)
- 6) 試験からの病理学的データの収集及び解釈を担当した試験施設名及び検査担当科学者の氏名
- 7) 発生率データを作成するために合計された可能性のある腫瘍の種類の記事

背景データは、試験ごとに絶対値や相対値など評価に役立つ値が示されるべきである。集計値には、値の範囲、平均値、中央値及び標準偏差を含むべきである。げっ歯類の稀な腫瘍の評価には文献データが適用できる場合もある。肝臓、下垂体及び副腎の腫瘍のように通常みられる腫瘍の場合は、発生率の範囲が広いために、投与群も含めた腫瘍の発生率が背景データの範囲内に収まることが多く、範囲だけでは評価が困難であり平均や分布が有用である。被験物質と類似の構造または作用機序を示す他の化学物質に関する組織学的データも有用である。もし、背景データの範囲内だが腫瘍発生数の軽度な増加が被験物質の標的臓器でみられた場合、前腫瘍性病変やその標的臓器への連続的な障害の有無について注意深く評価する必要がある。

(胎児の異常) 対照群のデータは試験ごとに必要である。これに対し、複数の試験から得られた陰性(溶媒) 対照群のデータで構成される背景データは、日常的に必要とされるものではないが、試験の所見を解釈するために有用で適切と考えられている。投与群のデータと比較する場合、背景データよりも当該試験の対照群のデータが優先されるべきである。胎児の異常に関する背景データも長期試験の背景データで上述した基準 1) から 6) に加え、7) 奇形や変異の評価基準が示される必要がある。

番号	2-07
著者名	OECD
文献名	Guidance Document 116 on the Conduct and Design of Chronic Toxicity and Carcinogenicity Studies, Supporting Test Guidelines 451,452 and 453--2nd Edition. ENV/JM/MONO(2011)47
巻, ページ, 発行年	2012
分類 (①~④)	②
参照箇所の該当ページ	4.4 (115, 116)

OECD TGs における統計学的手法の記載

- OECD TGs 451, 452, 453 の共通項 (275)
 - 数値結果は適切で一般的に許容される統計的方法で評価されるべき
 - 統計的手法及び分析すべきデータについては、試験計画の段階で選択されるべき
 - 必要に応じて、生存率で調整する規定を設けるべき
- TG451 及び 453 の共通項 (276)
 - 実験計画と目的を考慮し、最も適切な統計的方法は、試験開始前に確立されるべき
 - 生存率の補正, 生存期間に対する累積腫瘍リスクの分析, 腫瘍発生時期の分析, 及び1つまたは複数の群が途中で終了した場合の解析を統計手法に含めるか考慮する
 - 適切な統計分析に関するガイダンス及び国際的に認められた統計手法の主な参照事項は、OECD の試験ガイドラインの公開サイトで入手可能な慢性毒性及び発がん性試験の設計及び実施に関するガイダンス文書, ならびに慢性毒性及び発がん性試験の分析及び評価に関するガイダンス文書 No. 35 に記載されている
- OECD TGs 452 (276)
 - 生存率の補正及び1つまたは複数の群が途中で終了した場合の解析を統計手法に含めるかを考慮する

EPA 発がん性物質リスクアセスメントのガイドラインにおける統計学的手法の記載

- 長期試験の統計解析は、腫瘍の種類ごとに個別に行うべき
 - 同じ細胞由来の良性病変と悪性病変の発生率は、それぞれ別々に検討された後、科学的に説明可能な場合には組み合わせて検討される (McConnell ら, 1986 年)
 - 傾向検定及び pairwise 比較検定は、腫瘍発生率の明らかな増加が、被験物質投与に関連した影響というよりはむしろ偶発性であることを証明するために推奨される検定である
 - Cochran-Armitage 検定 (Snedecor & Cochran, 1967 年) のような傾向検定は、すべての投与群の結果が投与量の増加に伴って増加するかどうかを問うものである
 - Fisher exact 検定 (Fisher, 1950 年) のような pairwise 比較検定では、ある1つの投与群の発生率が対照群よりも増加しているかどうかを問うものである
 - 統計的に有意な差とは、発生率の増加が偶発性であることを示す p が 0.05 未満のものである
- どちらの種類の場合でも、結果が偶然によるものであるという仮説を棄却するには、有意性があれば

十分である。統計的に有意な反応は、生物学的に有意である場合もあれば、そうでない場合もあり、その逆もある。有意水準の選択は、偽陽性と偽陰性のリスクの間のトレードオフに基づいている。有意水準が5%以上または以下（最も一般的な有意水準）の場合、他の科学的情報を確認できるかどうかを検討する。評価が単純な5%レベルから逸脱する場合は、リスク判定の際に強調する必要がある。両側検定または片側検定を使用することは可能であるが、いずれの場合も根拠を示すことが必要である。

番号	2-08
著者名	Mann PC, et al
文献名	International Harmonization of Toxicologic Pathology nomenclature: An Overview and Review of Basic Principles
巻, ページ, 発行年	Toxicol Pathol. 2012; 40(4 Suppl): 7S-13S
分類 (①～④)	②
参照箇所の該当ページ	9S-11S, Table 1

グレーディングシステム/程度評価

- ・病変の程度は、変化の度合い（その所見に付随する変化がどれくらい存在するか）、分布（局所～びまん性）、及び実際の重症度の組み合わせを反映して選択されるべきである。
- ・グレーディングシステムは、研究施設により様々であり、病変の分布、病期、及び病変の範囲をどのように組み入れるかで異なる場合があるため、普遍的なグレーディングシステムを構築することは困難である。
- ・最も一般的に使用されているグレーディングシステムは、4～5段階の明確な半定量的評価法（表）である。
- ・被験物質の影響及び用量相関性の判断は、統計学的分析よりも担当病理医が実施することが適切である。
- ・同一化合物に関する試験内又は試験間での程度付の整合性は、病理学的評価の品質保持のため重要であり、試験内での程度付の不整合性又は「ドリフト (drift)」(診断基準が経時的に変化する傾向)は、被験物質の影響を検出する感度が損なわれるか、または実際には被験物質の影響がないにもかかわらず、それがあがる（偽陽性）ことにつながる可能性がある。
- ・投与に関連した影響や用量反応性を伝えるために、毒性病理医は試験内のスコアの基準を変更したり、「分割 (split)」したりすることがあるが、この場合は病理報告書の方法のセクションでそのプロセスと基準を明記すべきである。
- ・Peer review において、reviewer は、試験内でのグレーディングの一貫性を評価すべきである。

程度のない病変

- ・病変によっては、付加的な情報が得られないことから、程度付けを行わないことがある。
- ・これらの病変は通常、重症度の評価ではなく「存在する (P)」として記録される。
- ・例としては、腫瘍、嚢胞、自己融解、先天性異常などが挙げられる。
- ・一方で、病理データ入力システムの中には、すべての病変について等級を入力することを要求するものもある。

表1 4点採点システムの説明

数値スコア	説明	定義
0	正常の範囲内	試験の条件下で、関係する動物の週齢、性別、系統を考慮して、正常であると考えられる組織。他の状況下では正常からの逸脱と考えられる変化が存在する場合がある。
1	極軽度	正常範囲内であると考えられるものからわずかに超えた変化。
2	軽度	一般的に、病変は容易に同定できるが、限局的である。
3	中等度	病変は顕著であるが、さらに重症する可能性がある。
4	高度	変化の程度は可能な限り最大である（臓器の大部分を占めている）。

番号	2-09
著者名	Elmore SA, Peddada SD
文献名	Points to Consider on the Statistical Analysis of Rodent Cancer Bioassay Data When Incorporating Historical Control Data
巻, ページ, 発行年	Toxicol Pathol. 2009; 37(5): 672-676
分類 (①~④)	②
参照箇所の該当ページ	Abstract

概要

- ・被験物質の投与の影響を判断する最も関連性の高い対照データは、同時対照群のデータではあるが、特定の状況においては背景データ Historical control data (HCD) の評価も有用である。これには、自然発生腫瘍、希少腫瘍、一般的な腫瘍、発生数が激しく変動する腫瘍、同時対照群と比較して発生数が増加している腫瘍、または対照動物における腫瘍発生数の予期せぬ増減などの評価が含まれる。
- ・HCD は化合物の潜在的な発がん性を評価する際に、「根拠の重み」付けを補足できる多くの情報源の1つとして使用されるべきである。考慮すべき他の情報としては、類似の細胞由来の他の病変、体重、生存率、腫瘍の発生時期、その腫瘍が両種または雌雄に発生している場合、用量相関性がある場合、または両側性臓器の両側に発生している場合である。
- ・腫瘍発生数を HCD と比較する方法は2つあり、一方はデータの探索的（非公式）分析（exploratory analysis）、他方はより正式な統計的アプローチ（more formal statistical approach）分析である。
- ・箱ひげ図は、HCD の分布に関して現在のデータを説明する探索的なツールとして用いる。平均値、標準偏差、範囲などの標準的な方法ではわからない、潜在的な異常値を特定することができる。
- ・ここで述べられている HCD の統計解析に関する様々なオプションは、必ずしも標準的な手法を示すものではなく、これら以外にも種々の統計手法が存在している。

番号	2-19
著者名	OECD
文献名	Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies
巻, ページ, 発行年	2002
分類 (①～④)	②
参照箇所の該当ページ	27, 28, 36, Appendix V

主要試験項目の分析と評価

試験項目の分析と評価において、観察された被験物質の変化のすべてが必ずしも毒性作用ではないことを念頭に置くべきである。肝酵素誘導による肝肥大のような適応反応や感染症の治療に使用された抗生物質投与による白血球減少のような薬理作用などが例として挙げられる。

パラメータの中には加齢に伴って、計時的に値が変動するものがあるため、週齢の一致した対照群と比較すべきである。背景データを用いる場合でも背景データとの統計学的な比較は適切ではない。老齢げっ歯類に観察される慢性進行性糸球体腎炎やアミロイドーシスなどのような加齢性病変については、被験物質投与に影響を受ける可能性があるとしてされている。

背景データは、当該試験の対照群に異常がないかの確認にも有用である。対照群の背景データからの逸脱がみられる場合には、報告書においてその解析結果に対して何らかの考察が必要である。背景データとの間に一貫して差が認められる場合には、その動物の由来を含めて検討すべきである。

背景データを利用する場合には、以下の条件が必要である。

- ・試験から得られたデータは同じ試験実施場所で検査されたものであること。
- ・当該試験の動物は同じブリーダーからの同じ系統、年齢及び性別であること。
- ・当該試験の前後2～3年以内に実施されたものであること。
- ・試験条件（絶食・非絶食の条件、試験パラメータの想定方法、病理組織学的な判断基準、最終と殺時の年齢）なども明らかにする。

溶媒を用いた場合には、溶媒対照群を投与しない対照群も設定することで溶媒の影響の有無を知ることができる。また、動物に対する取扱い方法の違いによっても検査値に影響する可能性がある（例えば、マウスにおいて保定による血清ALT上昇が報告されており、対照群も同様に処置する必要がある）

基本的な検査項目（体重増加、食物消費量、変換効率など）は多くの場合、低用量で影響を受ける最も重要なエンドポイントである。評価者は、さらに、臓器重量、病理組織学的検査結果、血液生化学的検査値や神経毒性に関する項目も含め、さらに特定の生物学的または毒物学的影響に対する用量反応関係を慎重に考慮して評価に重要なエンドポイントを判断すべきである。

統計解析の方法

毒性試験の目的は生物学的に重要な反応を示すことであることを念頭に置く必要があることから統計結果を評価に用いる場合には以下の点に注意する。

- (1) 統計は質の悪い検査データを良くすることはできない。
- (2) 統計的有意性は生物学的な意義を意味するものではない。

- (3) 生物学的に意義のある影響であっても、統計的には有意でない場合がある。
- (4) 統計的有意性がないことが安全性の証明にはならない。

付録Vには、一般的に使用される統計的検定がいくつか記載されている。

なお、統計解析において以下の場合にはその旨を報告書に記載し、必要に応じて、データの再分析などの他の措置をとる必要がある。

- ・統計的検定が使用されていない場合、
- ・不適切な検定が使用されていると思われる場合、
- ・一般的に使用されていない検定が使用されている場合には、その旨を報告書に記載する。

付録 V 適用可能な統計的検定の一覧

非正規性の検定

- － カイ二乗
- － Kolmogorov-Smirnov
- － Shapiro-Wilk

分散の同質性の検定

- － バートレットの検定
- － Levene の検定

正規分布データの仮定

1. 全体的な検定

- － 分散分析 (ANOVA, 固定効果モデル, Model 1 ANOVA が最も一般的だが、他のモデルもあり得る)
- － 共分散分析 (ANCOVA)
- － Pearson の相関係数 (Pearson's Correlation Coefficient)
- － 用量効果の傾向を調べるための線形回帰(従属データである2つのパラメータの関係をテストする。回帰スロープの有意性を検証する。例：男女間の反応は似ているか?)

2. ペアワイズ比較

- － Duncan の多重範囲検定
- － Dunnett の t-検定 (コントロールと他のグループの平均値を比較する)
- － Scheffe の検定 (多重比較, Newman-Keuls の多重範囲検定よりも検出力が弱い)
- － Williams の t-テスト
- － Student の t-テスト
- － Fisher の最小有意差 (LSD) 検定

ノンパラメトリック手続き (パーセント値, 順位など)

- Kendall の順位相関係数
- Mann-Whitney U-Test (t-test に類するもの)
- Wilcoxon signed-rank 検定 (ペア, マッチドペアデータ)
- Kruskal-Wallis ANOVA
- 分布によらない多重比較検定 (例: Dunn の検定, Shirley の検定)
- Jonckheere の検定

定量的データ (死亡率, 病理所見など)

- Fisher の正確検定
- R × C カイ二乗検定
- Litchfield & Wilcoxon - ED50 などの信頼限界

多変量解析法

- ホテリングの T^2 法
- 多変量分散分析 (MANOVA)

海外のリスク評価機関における評価結果等に関する調査ー

3. (4) 残留農薬のリスク評価にあたっての毒性試験結果の解釈に係る知見の収集

③ 血液学的検査、血液生化学的検査及び尿検査に対する基本的な考え方及びこれらの検査において統計学的に得られる正常範囲

番号	著者名	文献名	雑誌：巻：ページ	参照箇所の該当ページ
1	World Health Organization (WHO)	Guidance Document for WHO Monographers and Reviewers	2015	Part 3: 1.7 (38, 39)
5	Hall AP, et al	Liver Hypertrophy: A Review of Adaptive (Adverse and Non-adverse) Changes--Conclusions from the 3rd International ESTP Expert Workshop	Toxicol Pathol. 2012; 40(7):971-94	976
6	Yokoyama Y, Ono A, Yoshida M, Mastumoto K, Saito M	Toxicological Significance of Increased Serum Alkaline Phosphatase Activity in Dog Studies of Pesticides: Analysis of Toxicological Data Evaluated in Japan	Regul Toxicol Pharmacol. 2019; 109:104482	Abstract
10	Friedrichs KR, Harr KE, Freeman KP, Szladvovits B, Walton RM, Barnhart KF, Blanco-Chavez j	ASVCP Reference Interval Guidelines: Determination of de novo Reference Intervals in Veterinary Species and Other Related Topics	Vet Clin Pathol. 2012; 41(4):441-453	Abstract, 441, 446-447, Table 4

番号	3-01
著者名	World Health Organization (WHO)
文献名	Guidance Document for WHO Monographers and Reviewers
巻, ページ, 発行年	2015
分類 (①～④)	③
参照箇所の該当ページ	Part 3: 1.7 (38, 39)

概要

血液学, 血液生化学及び尿検査の毒性評価では, 関連するパラメータが一貫して同じ方向 (増加あるいは減少) に向かっているか, 用量と関連しているか (又は, 「用量依存性であるか」) を判断することが重要である。血液学, 血液生化学及び尿検査の背景データとしては, 過去 10 年以内のデータや同一試験施設で実施された毒性試験と同時期のものが適切である。背景データは, 統計的範囲 (±1 標準偏差) で表すことが望ましい。年齢, 系統, 取り扱いや給餌の技術, 測定に使用する機器など, 多くの要因が背景データに影響を与える可能性がある。イヌでは, 個体値のばらつきが大きいので, 平均値よりも投与開始前を含めた試験期間中の個々の値の変化の方が有用である。また, 構造的に類似した化学物質または類似の作用機序を示す化学物質の血液学, 血液生化学及び尿検査パラメータに関する情報も有用である。尿検査結果は, 特に全身の栄養状態に影響するような高い濃度で投与する場合に, 常に他のエンドポイントと併せて解釈する必要がある。

毒性と考えられない影響には次のようなものがある。(血液検査) 1) 総白血球数の変化を伴わないリンパ球と好中球の数の変化, 2) 白血球の中のマイナーな細胞集団 (好塩基球, 好酸球, 単球) の数の 3 倍以内の変化, 3) 赤血球数, ヘマトクリット値またはヘモグロビン濃度の 10% 以内の増加または減少, 白血球数及び血小板数の 20% 以内の増加または減少, 4) 高用量で肝毒性によって二次的に起こった軽度な貧血 (血液生化学検査) 1) 組織・臓器に由来する酵素の僅かな増減 (AST, ALT, ALP, γ -GTP, CK 等) : 50% 以上で毒性指標を考慮, 2) 若齢動物の γ -GTP 活性を対照群との百分比で表した場合 (誤った解釈につながる), 3) 総タンパク濃度またはアルブミン濃度の変化を伴わないアルブミン/グロブリン比, 4) 10% 以下の総タンパク濃度またはグルコース濃度 (尿検査) 1) 排泄される酸性または塩基性化合物に関連する pH の変化, 2) 水分消費量または体液減少 (嘔吐または下痢など) の変化に関連する尿量及び/または比重の変化。

軽度ではあるが毒性学的に有意な変化には次のようなものがある。(血液検査) 1) 平均赤血球容積の 5% 以上の増加は赤血球の形態学的異常を示す可能性がある, 2) ヘモグロビン濃度が約 10% 以上低下した場合は貧血の指標となる, 3) メトヘモグロビンは正常でも存在するが, 投与群において, イヌで 5% 以上, ラットで 1.5% 以上の上昇は毒性と考えられる, 4) スルファメトヘモグロビンは正常では存在しないため, 検出されれば異常である。(血液生化学検査) IgG 濃度の低下は免疫抑制の指標となる。

(尿検査) 1) 細胞 (血液細胞, 腎臓上皮細胞または膀胱上皮細胞) 数の増加は腎臓・尿路系の障害の指標となる, 2) 尿中の血液は, 全身の出血や腎臓, 尿路, 雄性生殖器系における局所的な障害を示すことがある。

番号	3-05
著者名	Hall AP, et al
文献名	Liver Hypertrophy: A Review of Adaptive (Adverse and Non-adverse) Changes-- Conclusions from the 3rd International ESTP Expert Workshop
巻,ページ,発行年	Toxicol Pathol. 2012; 40(7): 971-94
分類 (①～④)	③
参照箇所の該当ページ	976

肝細胞 γ GT (γ -glutamyltranspeptidase) の誘導は、フェノバルビタールナトリウムを含むいくつかの化学物質でも報告されており、フェノバルビタールナトリウムは肝重量の増加を引き起こし、肝臓での代謝亢進を誘導することが知られている。ワークショップで発表された事例研究では、ラットを用いた反復投与 (7日間) の経口毒性試験において、肝 γ GT のメッセンジャーRNA (mRNA) が誘導されることが示され、小葉中心性肝細胞肥大、肝重量の増加、CYPs (主に 2B2 と 3A3) の増加、第2相酵素活性 (主に GSTA3 と UGT1A6) の増加が観察された。しかし、これは循環血中の γ GT 活性の増加にのみ影響し、肝 γ GT mRNA の顕著な誘導 (50 倍以上) が観察された。フェノバルビタールを5日間投与したラットでは、肝 γ GT 活性は増加したが、血清 γ GT 活性は増加しなかったことから、実験動物の肝胆道変化の指標としての血清由来 γ GT は、肝 γ GT の測定と比較して相対的に感度が低いことが示唆された (Goldberg et al. 1981)

しかし、肝内胆管の部分的な閉塞による胆汁閉塞も同様に γ GT 酵素活性の変化を引き起こす可能性がある。したがって、 γ GT の増加を酵素誘導によるものとする判断は、肝重量の増加、生物学的に意義のあるトランスアミナーゼ活性の増加がないこと、及び病理組織学的に有害な事象を示す変化がないことを十分に考慮した上で、証拠の重み付けに基づいて行われるべきである。

γ GT と ALP は、ラット、マウス、イヌにおいて異なる組織分布を有する (Clampitt and Hart 1978; Keller 1981)。例えば、イヌでは、肝型が血中の ALP 活性の主体となるが、ラットでは比較的活性が低く、このため、この種では診断感度が低い。さらに、レギュラトリーな毒性試験に一般的に用いられる幼若な動物では、血清 ALP 活性は主に骨のアイソザイムに由来し、肝臓には由来しない (Hoffmann et al. 1994)。したがって、肝臓酵素の誘導に加えて、他の臓器での変化も考慮した評価をする必要があるかもしれない。

前述の化学物質による酵素の誘導もまた、種特異性があると考えられる：イヌでは、コルチコイド誘発性 ALP アイソザイム活性がしばしば増加するのに対し (Gaskill ら 2005)、マウスでは肝 ALP が増加する (Kawasaki, Mataka, and Takano 1994)。ラットでは、肝酵素誘導剤に反応して肝 γ GT が誘導される (Gallagher ら 1998 ; Satoh ら 1982)。妊娠6日目 (GD6) から GD20 までの妊娠ラットでは、酵素誘導剤を投与すると、 γ GT 活性ではなく ALP 活性が上昇する (Strauss, pers.comm. 2011)。マウスでは、ラットとは対照的に、肝細胞毒性の病理組織学的徴候を伴わない肝酵素誘導剤を投与した場合には、ALT が増加することが多い (Strauss, pers. comm. 2011)。

また、市販のヒト用 γ GT キットの検出限界は約 3 U/L であることに注意が必要である。(37°C で 50 nkat/L, 例として γ GT, Szasz mod., 2010-06, V7 : ロシユ取扱説明書)。正常なラット血清 γ GT 活性は、一般的にこの検出限界以下であるので、中等度の γ GT 増加は、これらの検査では検出できない。

番号	3-06
著者名	Yokoyama Y, et al
文献名	Toxicological Significance of Increased Serum Alkaline Phosphatase Activity in Dog Studies of Pesticides: Analysis of Toxicological Data Evaluated in Japan
巻,ページ,発行年	Regul Toxicol Pharmacol. 2019; 109: 104482
分類 (①～④)	③
参照箇所の該当ページ	Abstract

概要

血清アルカリホスファターゼ (ALP) 活性の上昇は、ヒトや実験動物における肝障害の指標となる。実際には、イヌを用いた農薬の毒性試験において、他の肝障害の変化を伴わないALPの上昇がしばしば認められる。本研究では、食品安全委員会がリスク評価の過程で毒性評価報告書として評価した206種の農薬について、ALP上昇の毒性学的意義を解析した。その結果、ラット(36/206)よりもイヌ(108/206)の方がALPの上昇が認められる試験の割合が高かった。108農薬中87農薬では、イヌで肝毒性を伴うALP上昇が認められた。しかし、ALPの増加には、特定の種類の肝毒性との特異的な関連は認められず、肝毒性のマーカーとしては高感度(または鋭敏)とはいえなかった。肝毒性を示した87種の農薬の約50%は、肝肥大も誘発した。残りの21種類の農薬では、肝重量の増加や肝肥大以外に肝毒性の変化はみられなかった。これら21種類の農薬のほとんどは、齧歯類におけるフェノバルビタール様肝代謝酵素誘導剤であった。これらの結果から、肝毒性所見がない場合、ALPの上昇はイヌにおける肝毒性の指標にはならないことが示唆された。

解析の結果、イヌにおけるALPの上昇についてその毒性学的意義を決定する要因が明らかとなった。まずALPの上昇が肝毒性や肝肥大といった異常所見を伴っているかについて確認する。肝毒性に加え肝肥大が認められる場合、ALPの増加が認められる用量と肝毒性が認められる用量を比較し、ALP増加が認められる用量が肝毒性のそれに比べ同等あるいはより高い場合、ALPの上昇は有害影響であり、肝毒性のNOAELには影響を与えない。ALP増加がみられる用量が、肝毒性がみられる用量より低い場合、さらにその用量について肝肥大を示す用量との比較を行う。ALP増加がみられる用量と肝肥大がみられる用量が同等であった場合、ALPの上昇は有害影響とはならず、肝毒性のNOAELに影響を与える。同等でなかった場合には、期間の延長により毒性が増強されることを示す証拠があるかを考慮し、証拠が認められる場合ALPの上昇は有害影響であり、肝毒性のNOAELには影響を与え、認められない場合はALPの上昇は有害影響とはならず、肝毒性のNOAELに影響を与える。肝毒性のみ認められる場合、ALPの増加が認められる用量において肝毒性が認められる用量に比べ同等か高いこと、あるいは期間の延長により毒性が増強された証拠の有る、といった条件のいずれかに当てはまる場合ALPの上昇は有害影響であり、肝毒性のNOAELには影響を与えず、何れにも当てはまらないのであれば、ALPの上昇は有害影響ではなく、肝毒性のNOAELに影響を与える。肝肥大のみ伴っている場合、肝代謝酵素の誘導や肝肥大を示す証拠があればALPの上昇は有害影響ではなく、肝毒性のNOAELに影響を与える。そういった証拠がない場合、期間の延長によりさらに低い用量での毒性の増強を示す証拠の有無を確認し、有る場合はALPの増加はおそらく有害影響であり、肝毒性のNOAELに影響を与えることになる一方、無い場合はALPの上昇は有害影響ではなく、肝毒性のNOAELに影響を

える。他の肝臓所見が見られない場合も同様の考察が可能であり、この解析はイヌにおける ALP の毒性学的意義の新たな解釈を提供し、農薬の毒性学的評価に寄与する可能性がある。

番号	3-10
著者名	Friedrichs KR, et al
文献名	ASVCP Reference Interval Guidelines: Determination of de novo Reference Intervals in Veterinary Species and Other Related Topics
巻, ページ, 発行年	Vet Clin Pathol. 2012; 41(4): 441-453
分類 (①～④)	③
参照箇所の該当ページ	Abstract, 441, 446-447, Table 4

概要

基準間隔Reference intervals (RI) は、検査室での診断検査や臨床上の意思決定に不可欠な要素であり、同等の個人の健康な集団からの基準値Reference values (RV) の推定分布を表す。診断の追求や治療開始の意思決定は、しばしばRIに含まれない値に基づいて行われるため、RV値の収集と分析には細心の注意を払う必要がある。

慣例上、RIはRVの中心95%を含み、基準値の上限と下限で区切られる。基準限界値を決定するために選択された統計的方法は、RVの数と分布に基づく(表1)。ノンパラメトリック法は、少なくとも120個のRVが利用可能な場合に推奨される。2.5分の1と97.5分の1のフラクタイルfractilesがそれぞれ基準値の下限と上限として機能する。利用可能なサンプル数が120未満の場合、90%信頼区間を決定するためには、代替の方法(ブートストラップ等)が必要となる。39個は、95%ノンパラメトリックRIを決定できる最小サンプル数であるが、この場合、極端な値が下限値と上限値となる。ノンパラメトリック法を小さなサンプルサイズで使用しなければならない場合は、両極端の潜在的な外れ値をトリミングできるように十分なサンプルを収集する必要がある。あるいは、ロバスト法またはパラメトリック法を使用すべきである。ロバスト法は、 ≥ 40 及び ≤ 120 の参照サンプルが利用可能な場合に推奨され、データの位置と広がりやを推定するために反復プロセスを利用する。またガウス分布を必要としないが、データが対称的に分布している場合、より良いパフォーマンスを発揮する。ブートストラップ法は90%信頼区間を決定するために使用されるべきである。

データがガウス分布を持つか、またはガウス分布に変換できる場合、 ≥ 40 及び ≤ 120 の基準サンプルが利用可能な場合、パラメトリック法を使用することができる。パラメトリック法は、値の中心95%よりもわずかに多くの値を包含し、それぞれ平均+2SD及び平均-2SDで上限及び下限の基準限界を確立する。パラメトリック法は、基準値の90%信頼区間を決定するために使用されるべきである。

獣医学では、特殊種や野生で捕獲された種、新生児など、基準サンプルの収集数が限られる場合がある。20～40サンプルの基準サンプルの場合、RIはロバストまたはパラメトリックな方法で計算する必要がある。小さなサンプルサイズに内在する不確実性を強調するために、90%信頼区間を計算すべきである。さらに、情報に基づいた臨床的意思決定を可能にするために、ヒストグラム、平均値や中央値、最小値及び最大値を報告すべきである。あるいは、ヒストグラムと共にすべての値を列記した表の提供に代替され得る。

10人以上20人未満の基準サンプルが入手可能な場合は、ヒストグラム、平均値または中央値とともに昇順値の表を報告すべきであるが、少ないサンプル数に基づく限界値の不確実性を考慮して、RIを決定すべきではない。参照個体が非常に少ない場合には、被験者ベースのRIを考慮すべきである。一般に、利用可能な

参照個体が 40 人未満の場合、適切な参照被験者の選択に細心の注意を払い、標準化された収集技術と十分に管理された分析方法を遵守して、外れ値のないサンプルを収集することに重点を置くべきである。外れ値の存在を評価することは特に重要で、単一の外れ値の存在は推定基準値に強い影響を与えるためである。外れ値が識別された場合は、代替サンプルを収集するためにあらゆる努力をすべきである。

表 1. 参考サンプルサイズと分布に基づいた RI の設定のための推奨手順

サンプルサイズ	データ分布 (生または変換)	統計的手法
≥ 120	該当なし	参照限界の 90%信頼区間 (CI) を持つノンパラメトリック
$40 \leq x < 120$	ガウス	参照限界の 90%CI でロバスト 基準値の 90%CI を持つパラメトリック
	非ガウス	基準限界の 90%CI でロバスト (望ましい) ノンパラメトリック
$20 \leq x < 40$	ガウス	参照限界の 90%CI を持つパラメトリック†
	非ガウス	参照限界の 90%CI でロバスト†
$10 \leq x < 20$	該当なし	基準間隔を計算しない†
< 10	該当なし	参照値を報告しない

* 120 個未満の参照サンプルでは、ノンパラメトリックに 90%信頼区間を決定できない；ブートストラップなどの代替方法が必要となる。

† ヒストグラム、平均値または中央値、最小値及び最大値を含む；代わりに、ヒストグラムと一緒にすべての参照値の表を提供することができる。

海外のリスク評価機関における評価結果等に関する調査－

3. (4) 残留農薬のリスク評価にあたっての毒性試験結果の解釈に係る知見の収集

④ 体重、摂餌量及び臓器重量に対する基本的な考え方及びこれらの重量について統計学的に得られる正常範囲

番号	著者名	文献名	雑誌：巻：ページ	参照箇所の該当ページ
1	World Health Organization (WHO)	Guidance Document for WHO Monographers and Reviewers	2015	Part 3: 1.8 (40, 41)
11	Everds NE, Snyder PW, Bailey KL, Bolon B, Creasy DM, Foley GL, Rosol TJ, Sellers T	Interpreting Stress Responses during Routine Toxicity Studies: A Review of the Biology, Impact, and Assessment	Toxicol Pathol. 2013; 41(4): 560-614	Abstract, 561-562, 563, 592, 595
12	Levin S, Semler D, Ruben Z	Effects of Two Weeks of Feed Restriction on Some Common Toxicologic Parameters in Sprague-Dawley Rats	Toxicol Pathol. 1993;21(1):1-14.	Abstract , 5, 13
13	Michael B, M, Yano B, Sellers RS, Perry R, Morton D, Roome N, Johnson JK, Schafer K	Evaluation of Organ Weights for Rodent and Non-rodent Toxicity Studies: A Review of Regulatory Guidelines and a Survey of Current Practices	Toxicol Pathol.2007; 35: 742-750	742-743, 749, 750
14	Marino DJ (2012b)	Age-Specific Absolute and Relative Organ Weight Distributions for Fischer 344 Rats	J Toxicol Environ Health A. 2012; 75(24): 1484-1516	Abstract
15	Marino DJ (2012a)	Age-Specific Absolute and Relative Organ Weight Distributions for B6C3F1 Mice	J Toxicol Environ Health A. 2012; 75(2): 76-99	Abstract
16	Dorso L, Chanut F, Howroyd P, Burnett R	Variability in Weight and Histological Appearance of the Prostate of Beagle Dogs Used in Toxicology Studies	Toxicol Pathol. 2008; 36: 917-925	Abstract, 923, Figure 13
17	Bailey SA, Zidell RH, Perry RW	Relationships Between Organ Weight and Body/Brain Weight in the Rat: What Is the Best Analytical Endpoint?	Toxicol Pathol. 2004; 32: 448-466	Abstract, 461, 466, Table 3
18	Sellers RS, Morton D, Michael B, Room N, Johnson JK, yano BI, oerry R, Schafer K	Society of Toxicologic Pathology position paper: organ weight recommendations for toxicology studies	Toxicol Pathol. 2007; 35: 751-755	752, Appendix 1
19	OECD	Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies	2002	30-31, 33-34

番号	4-01
著者名	World Health Organization (WHO)
文献名	Guidance Document for WHO Monographers and Reviewers
巻, ページ, 発行年	2015
分類 (①~④)	④
参照箇所の該当ページ	Part 3: 1.8 (40, 41)

概要

正常な生物学的変動内での影響に関する一般的な結論

主に体重または体重増加, 臓器重量変化, 血液学的・生化学的・尿検査パラメータに関する影響である場合, 文献で公表されている正常範囲の値は, 当該試験で使われた方法と同等であることを保証するために, それらが得られた方法が明記されている場合にのみ使用すべきである。それらの影響のパターンを考慮すると, 毒性学的影響の概算の変化の大きさは, 1) 10%の体重または体重増加, 2) 20%のアセチルコリンエステラーゼ阻害, 3) 臓器重量の変化がある。

臓器重量の変化は, 毒性試験で観察された被験物質投与に関連した影響のうち, しばしば最も感度の高い指標となる。臓器重量の解析は, 絶対重量, 相対重量の両方を基に行うべきである。臓器重量の変化は, 絶対値及び相対値の両方とも, 対照群と投与群との差をパーセンテージで表したものである。体重変化が大きい場合, 相対重量は体重よりも脳重量との相対値で表す方が好ましい場合がある。臓器重量の変化は, 絶対重量, 相対重量, またはその両方が変化しているかどうかという観点で評価すべきである。その変化は, 適切な対照群の平均値と比較した百分比 (増加または減少) として表現すべきである。臓器重量の適切な評価には, 個々の動物の値と群の平均値の両方が必要である。ラットにおける標的臓器毒性を最も正確に検出しやすいエンドポイント (臓器重量, 臓器重量体重比, 臓器重量脳比) を決定するために実施された評価では, 臓器重量体重比は, 肝臓及び甲状腺, 臓器重量脳比は卵巣及び副腎への影響を評価するのにそれぞれ役立つ。脳, 心臓, 腎臓, 下垂体及び精巣の重量はいずれを選んでも適切に評価できないので, 共分散分析などの代替方法を利用すべきである。マウスにおいて投与に関連した臓器重量の変化を明らかにするためには, 肝臓, 腎臓, 肺及び心臓の絶対重量の加齢に伴う増加, 及び脳と精巣の重量が比較的安定であることを考慮しなければならない。マウスにおける, 絶対重量のばらつきの一般的な傾向として, 脳 < 精巣 < 腎臓 < 心臓 < 肝臓 < 肺 < 脾臓 < 胸腺と報告されている。ラットについては, 脳, 肝臓, 腎臓, 肺, 心臓, 甲状腺及び精巣の加齢による絶対重量の増加と相対重量の減少を考慮しなければならない。ラットにおける絶対重量のばらつきの一般的な傾向として, 脳 < 精巣 < 心臓 < 腎臓 < 肝臓 < 肺 < 胸腺 < 甲状腺と報告されている。

臓器重量データは, 剖検所見, 臨床検査結果及び病理組織所見とともに総合的に解釈しなければならない。対応する肉眼所見や組織所見を伴わない臓器重量の変化は注意して解釈する必要がある。臓器重量の変化が投与に関連したものであるかどうかを判断するためには以下の付加的な点を考慮すべきである。1) 用量相関性, 2) 統計学的有意差があるか, 3) 両性で起こっているか, 片性のみか, 4) 各投与群での発生率はどうか, 5) 対照群の範囲外か, 6) 臓器重量の変化が体重の変動に依存しているか, 7) 回復期間内で部分的または完全に回復するか, 8) 臨床検査結果, 剖検所見, 組織所見と関連があるか, である。

番号	4-11
著者名	Everds NE, et al
文献名	Interpreting Stress Responses during Routine Toxicity Studies: A Review of the Biology, Impact, and Assessment
巻, ページ, 発行年	Toxicol Pathol. 2013; 41(4): 560-614
分類 (①～④)	④
参照箇所の該当ページ	Abstract, 561-562, 563, 592, 595

概要

毒性試験の目的は、被験物質によって引き起こされた標的臓器の有害影響を特定することであり、高用量で観察される被験物質投与に直接的に起因する影響（一次的影響）は二次的にストレスをもたらすのに十分な程度と持続時間をもつ。被験物質の一次的影響は一般的に投与用量に比例するため、ストレスに起因する変化も用量に比例すると考えられる。このため、毒性試験における被験物質の一次的影響とストレスに関連した二次的影響とを区別することは困難である。ストレスが脳で感知されると、主に交感神経-副腎髄質系の活性化によるカテコールアミンの放出及び視床下部-下垂体-副腎系（HPA系）を介した血清グルココルチコイドの上昇が起こる。これらの神経内分泌シグナルにより、下流では総体重または体重増加量、摂餌量ならびに活動量の減少、臓器重量の変動（胸腺及び脾臓重量減少、副腎重量増加）、胸腺及び脾臓のリンパ球減少、循環白血球数の変動（リンパ球数及び好酸球数の減少を伴う好中球数の増加）、生殖機能の変化など、多臓器にわたり幅広い影響がもたらされる（表1）。これらの所見の一部またはすべてが発生した場合、それらはストレスに起因する可能性が高い。毒性試験でストレスを示す最も感度の高いパラメータは、循環白血球数、リンパ球サブセット、副腎及び胸腺重量であり、ストレスに対する感受性は動物種、ストレスの種類、動物の状態や個体差によって異なる。また、急性ストレスに関連した影響は臓器外観の変化ではなく、臨床病理パラメータの変化（例：ストレス性の白血球像）及び/または感受性の高い細胞集団の組織学的変化（例：グルココルチコイド誘発胸腺リンパ球のびまん性アポトーシス）として表れる。これに対し、慢性ストレスは体重及び臓器重量に影響を与える可能性がある（体重、胸腺、副生殖器及び脾臓重量の減少ならびに副腎重量の増加）。

毒性試験でみられる変化が被験物質ばく露に起因するものか、ストレスに起因するものかどうかを区別するためにはデータを総合的に評価する必要があり、以下に示す証拠の重み付けアプローチ（weight-of-evidence approach）が有用である。

- ・被験物質の標的臓器を考慮する：被験物質がストレスを受けやすい臓器に影響を及ぼす場合、その変化がどちらの影響によるものなのかを特定することは難しいため、低用量での被験物質の影響から判断材料を得られる可能性がある。

- ・ストレスによる影響を受けやすいパラメータや臓器におけるストレス反応の大きさを評価する：ストレス反応の大きさは与えられたストレスの大きさに見合ったものでなければならない（例：1ヵ月程度の軽度の反復ストレスを与えられた動物は副腎の重量及び組織に軽微から軽度の影響がみられる可能性がある）

- ・ストレスに関連した変化の用量反応と被験物質に関連した変化の用量反応との比較：ストレスに

関連した変化は一般的に被験物質の影響がみられない用量では起こりえない（例：中間用量及び高用量群で投与用量に関連した循環リンパ球及び胸腺リンパ球の減少，副腎重量の増加及び体重減少が認められ，低用量群では副腎重量の増加のみみられる試験では，低用量の変化は被験物質投与の直接的な影響である可能性がある）。

・他の解釈の可能性を考慮する：ストレスに起因する変化のほとんどは，別の解釈をもつ。データの解釈を行う前に，別の解釈とそれを支持または反論する証拠について十分に検討する必要がある（例：副腎皮質の肥大/過形成はストレスまたはグルココルチコイドの産生を抑制する副腎毒性物質が原因である可能性が考えられるが，他にストレスに関連した変化が認められない場合は副腎皮質ホルモン合成に対する影響である可能性がある）。

・証拠の重み付けの強さを評価する：一般的に感度の低いストレスの指標は，感度の高いストレスの指標が認められない場合には起こりえない（体重やリンパ系への影響等がなく生殖器系の影響のみがみられる場合，ストレスの影響ではない可能性がある）。

表 1. 毒性研究でよく評価される In-life 及び病理学的パラメータ：ストレスに対する典型的な反応

試験段階	影響を受けるシステム	パラメータ	起こり得るストレス影響 (影響がない場合もある)	
In-life	栄養摂取	体重または体重増加量	減少	
		摂餌量	減少	
臓器重量 ^a	循環血液細胞	好酸球数	減少	
		リンパ球数	増加または減少	
		好中球数	増加	
		免疫系	マクロファージの貪食作用	減少
	内分泌系	副腎	増加	
		免疫系	胸腺	減少
			脾臓	減少
	生殖器系	精巣 ^b	変化なし(ラット)または減少(マウス)	
		精巣上体	減少	
		精囊	減少	
前立腺		減少		
卵巢		減少		
子宮		減少		
生殖器系		精巣	変化なし(ラット)または変性(マウス)	
臓器組織学	消化器系	胃潰瘍	増加	
		リンパ系	胸腺細胞数	減少
	生殖器系	脾臓細胞数	減少	
		精巣	変化なし(ラット)または変性(マウス)	
		精巣上体，前立腺，精囊	萎縮の可能性あり	
	内分泌系	卵巢，子宮	機能低下	
		膾	萎縮，粘液産生	
		副腎皮質	肥大/過形成	

注：ストレスの影響を受けるパラメータは，動物種，性別，週齢及びストレス因子の種類と持続時間に依存する。種特異的反応については，該当部分を参照のこと。

a: 絶対重量，体重に対する相対量，及び/または脳重量に対する相対量

b: 絶対重量のみ

・免疫系とストレス

免疫系に対するストレスの生理学的及び形態学的影響は、血液、胸腺、脾臓、骨髄で見られる。リンパ系組織の中では、胸腺が最も感受性が高く、ストレスに対する反応は数時間以内に誘発される。グルココルチコイドは、リンパ球の再分布、細胞応答の変化、及びリンパ球の破壊を引き起こす可能性がある。臓器重量が約 20%以上減少した場合にリンパ球減少の組織学的所見と関連が認められる。未熟リンパ球はストレスの影響を最も受けやすいため、胸腺の皮質はリンパ球減少の結果、しばしばサイズが減少するが、骨髄や脾臓では胸腺に比べて成熟リンパ球が優勢であるため、これらの臓器でのストレス反応はそれほど明瞭ではない。ストレスに関連した変化は、他の二次リンパ系臓器ではあまり一貫性が認められない。

・胸腺

胸腺では、ストレスは皮質リンパ球の減少に起因する皮質のサイズの減少をもたらす。未熟な皮質リンパ球が最も影響を受ける。重度の急性ストレスは、多数の“tingible-body”マクロファージと細胞の破片（すなわち、starry sky 像）の出現を特徴とするびまん性のアポトーシスを引き起こす可能性がある。対照的に、軽度な慢性ストレスは皮質のサイズの明らかな減少をもたらすかもしれないが、アポトーシスの証拠は限られているか、全くない。胸腺リンパ球のアポトーシスはラットの拘束ストレスの2時間後に始まるが、胸腺重量は拘束後8時間まで変化しない。ラットの胸腺リンパ球数と重量は、毎日の強制水泳ストレスを21日間受けると減少する。マウスの胸腺重量は、ケージラックを短時間移動させた12時間後と24時間後にそれぞれ12%と15%減少した。胸腺に対する効果は可逆的であり得る。例えば、マウスにデキサメタゾンを単回投与すると、胸腺萎縮が生じ（細胞数及び増殖の減少によって示される）、これは処置後3日目までに最も顕著であるが、12日目までに完全に回復する。

・脾臓

毒性試験では、一般に脾臓に影響を及ぼすほどの十分に強いストレス因子は、リンパ球のアポトーシスの有無に関係なく、白脾髄領域の減少を伴う脾臓重量の減少をもたらす。ストレスに関連した変化は胸腺に比べて脾臓では一般的に一貫性がなく軽度である。1回の急性ストレスが脾臓重量に一過性の影響を与えることがマウスで報告されている。1日2時間のストレスに7日間連続でばく露されたラットでは、Tリンパ球領域の増殖が低下して白脾髄量が減少する。さらに、アポトーシスの増加によりBリンパ球領域が減少する。前述したように、循環Bリンパ球はストレス時にリンパ系器官に再分配し、その結果、ストレス下での脾臓のBリンパ球領域の可変的な変化に寄与している可能性がある。脾臓重量を増加させる被験物質の薬理的または毒性学的効果は、ストレスに関連した脾臓重量の減少を隠蔽する可能性がある。

番号	4-12
著者名	Levin S, et al
文献名	Effects of Two Weeks of Feed Restriction on Some Common Toxicologic Parameters in Sprague-Dawley Rats*
巻, ページ, 発行年	Toxicol Pathol. 1993;21(1): 1-14.
分類 (①～④)	④
参照箇所の該当ページ	Abstract, 5, 13

概要

本研究は味の悪い餌や食欲不振によって起こる短期の摂餌量減少が引き起こす変化を特定することを目的とし、4-5週齢の Sprague-Dawley SD ラット (Charles River) 各群雌雄 10 匹ずつの 4 グループに、2 週間にわたり異なる量の餌 [グループ 1: 自由摂取 (対照群), グループ 2: 前日の対照群の摂餌量の 75% (軽度食餌制限), グループ 3: 前日の対照群の摂餌量の 50% (中等度食餌制限), グループ 4: 前日の対照群の摂餌量の 25% (重度食餌制限)] を与え、血液学的検査, 血液生化学的検査, 臓器重量測定 (副腎, 脳, 心臓, 腎臓, 肝臓, 精巣及び胸腺) 及び組織学的検査 (臓器重量測定を行った臓器に加え, 骨及び骨髄 (胸骨), ならびに胃) を実施した。その結果, いずれの群にも動物の死亡は認められず, すべての食餌制限群において発育遅延以外の臨床症状は観察されなかった。試験終了時の最終体重は, 雄ではグループ 2, 3 及び 4 においてそれぞれ対照群と比較して 77%, 62%, 38% であり, 雌では 82%, 66%, 42% であった。血液学的検査では, 赤血球数はヘマトクリット値及び血色素量とともに食餌制限の程度に比例して増加し, 血小板数はすべての食餌制限群で減少した。白血球数はグループ 4 で著しく減少し, 特にリンパ球数の減少が目立った。血液生化学的検査では, グループ 4 の雌雄で血清ビリルビンの増加, 電解質異常及びコレステロールの減少 (雌のみ) がみられた。臓器重量では, 体重への著しい影響に伴い, グループ 4 のすべての臓器の絶対重量が減少していた。胸腺はグループ 4 の雌雄及びグループ 3 の雄で, 肝臓はすべての食餌制限群で絶対及び相対重量が減少した。その他の臓器の相対重量については, 変化しないか, もしくは増加していたが, これは減少した体重の大部分が脂肪, 筋肉及び水分の喪失によるものであったことを示唆していた。食餌制限に関連した組織学的所見は骨髄, 胸腺, 胃及び精巣で観察された [発生頻度は雌雄合算し (n=20), 病変の程度は軽微 (1), 軽度 (2), 中等度 (3), 高度 (4), 重度 (5) の 5 段階で評価した]。食餌制限により最も生物学的に重大な組織学的変化が観察されたのは骨髄であり, グループ 2 及び 3 では軽微から中等度の造血組織の減少 (骨髄腔の脂肪増加), グループ 4 では脂肪がみられず, 中等度から重度の造血及び間質成分の変性及び壊死が観察され, 造血組織は対照群の 10~25% であった。胸腺ではグループ 4 のすべての動物に退縮または萎縮が認められ, 雌でより強い影響がみられた (重度の萎縮は雄 1/10 及び雌 7/10 に観察された)。重度の胸腺萎縮では, ハッサル小体とわずかに残るリンパ球とともに泡沫状マクロファージが充満する間質しか残っておらず, 正常構造は失われていた。グループ 4 の残りの動物の胸腺には, マクロファージ (星空像を呈する可染体マクロファージ), 皮質のリンパ球減少, ハッサル小体の増加及び皮髄境界部の不明瞭が観察され, より程度の弱い萎縮像と考えられた。グループ 2 及び 3 では, 少数の動物に皮質における泡沫状マクロファージの軽微な増加が観察された。精巣の組織学的変化は, グループ 4 において 4 例にみられ, このうち 2 例では軽度な精上皮変性 (精上皮の脱落及び巨細胞の

形成)が、残りの2例では精子数の減少が観察された。胃では、剖検時にすべての食餌制限群で粘膜の局所的な赤色化もしくは血様内容物が認められ、組織学的に観察された軽度なびらんの発生頻度は食餌制限群で高かった(グループ 1-4: 1/20, 5/20, 2/20, 5/20)。副腎、脳、心臓、腎臓、または肝臓では、食餌制限に関連する組織変化は観察されなかった。

以上の結果より、生物学的に特に重要であると考えられる変化を以下に示す。

- ・ 摂餌量が対照群の75または50%の場合、胸骨骨髄における造血組織の減少及び脂肪の増加
- ・ 摂餌量が対照群の25%しかない場合、骨髄の壊死、胸腺萎縮(特に雌で強い)、骨髄及びリンパ組織へ影響を反映した好中球、リンパ球及び血小板減少、脱水の影響による血液濃縮及び血清電解質の異常、ならびに肝萎縮(絶対及び相対重量の減少)。その他の変化については、それほど重要ではないと考えられた。

(翻訳注: 肝臓の萎縮については相対重量の変化は統計学的に有意ではなく、組織学的な変化は観察されていないが、生体異物の代謝に影響を及ぼすかもしれないので無視出来ないとしている。)

表 IV. 臓器重量 (g) 及び臓器: 最終体重比; 算術平均, 標準偏差及び検査例数

	副腎		脳		心臓		腎臓		肝臓		胸腺		精巣		
	FBW (g)	WT	ORG:FBW ($\times 10^{-3}$)	WT	ORG:FBW ($\times 10^{-3}$)	WT	ORG:FBW ($\times 10^{-3}$)	WT	ORG:FBW ($\times 10^{-3}$)	WT	ORG:FBW ($\times 10^{-3}$)	WT	ORG:FBW ($\times 10^{-3}$)		
Males															
Group 1	213.4 8.92 10	0.04 0.015 10	1.89 0.736 10	1.9 0.10 10	8.9 0.64 10	0.9 0.06 10	4.4 0.28 10	2.1 0.23 10	9.9 0.98 10	7.8 0.54 10	3.7 0.18 10	0.58 0.138 10	2.71 0.657 10	2.24 0.066 10	10.51 0.413 10
Group 2	180.0*** 7.82 10	0.04 0.008 9	2.21 0.466 9	1.8* 0.07 9	10.1*** 0.45 9	0.8*** 0.04 9	4.3 0.29 9	1.6*** 0.07 9	9.1 0.42 9	6.1*** 0.41 9	3.4 0.15 9	0.57 0.157 9	3.18 0.888 9	2.21 0.167 9	12.29*** 0.634 9
Group 3	151.2*** 6.96 10	0.04 0.007 9	2.57* 0.474 9	1.8* 0.09 10	11.9*** 0.52 10	0.6*** 0.05 9	3.9 0.42 9	1.4*** 0.14 9	9.0 0.58 9	5.4*** 0.38 9	3.6 0.21 9	0.33*** 0.077 10	2.21** 0.495 10	2.22 0.192 10	14.69*** 0.915 10
Group 4	90.1*** 4.46 10	0.03** 0.005 8	3.21*** 0.580 8	1.7*** 0.06 10	18.7*** 1.31 10	0.4*** 0.05 10	4.9** 0.41 10	1.0*** 0.06 10	11.1*** 0.61 10	3.0*** 0.30 10	3.4 0.39 10	0.07*** 0.029 10	0.78*** 0.303 10	1.34*** 0.306 10	14.80*** 3.322 10
Females															
Group 1	144.4 16.36 10	0.05 0.010 10	3.62 0.741 10	1.8 0.10 10	12.2 1.05 10	0.7 0.13 10	5.0 0.54 10	1.4 0.16 10	10.0 0.48 10	5.3 0.73 10	3.7 0.26 10	0.48 0.123 10	3.33 0.679 10		
Group 2	131.5* 10.50 10	0.04 0.009 9	3.43 0.683 9	1.7 0.09 10	13.1* 0.78 10	0.6** 0.04 10	4.5 0.32 10	1.2** 0.10 10	9.4 0.51 10	4.8 0.42 10	3.7 0.13 10	0.44 0.089 10	3.35 0.723 10		
Group 3	108.7*** 6.06 10	0.04*** 0.006 10	3.29 0.483 10	1.7 0.09 10	15.3*** 0.64 10	0.5*** 0.08 10	4.6 0.68 10	1.1*** 0.10 10	10.0 0.68 10	3.7*** 0.40 10	3.4 0.27 10	0.36 0.133 10	3.35 1.227 10		
Group 4	64.7*** 5.85 10	0.02*** 0.007 9	3.65 1.072 9	1.6*** 0.06 10	24.3*** 1.43 10	0.3*** 0.06 10	5.2 0.58 10	0.8*** 0.12 10	12.4*** 1.04 10	2.3*** 0.21 10	3.6 0.23 10	0.03*** 0.024 10	0.50*** 0.320 10		

略語: FBW = 最終体重, ORG: FBW = 臓器: 最終体重比, WT = 重量

*, **, または*** 5, 1, または0.1%レベルでそれぞれ有意な傾向検定(フラグはそれぞれ平均値間の有意差を示す)

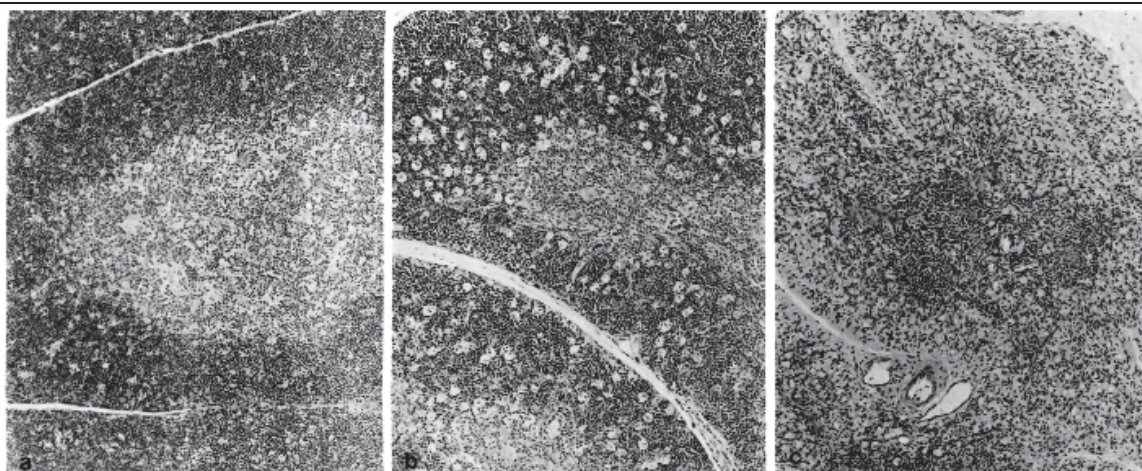


図 5- (a) 対照群のラット正常胸腺。(b) 軽度のリンパ球減少と皮質の星空像を示すグループ 4 (重度食餌制限) のラット胸腺。(c) 重度にリンパ球が減少し、胸腺構造が失われているグループ 4 (重度食餌制限) のラット胸腺。HE 染色。130 倍

番号	4-13
著者名	Michael B et.al. (2007)
文献名	Evaluation of Organ Weights for Rodent and Non-rodent Toxicity Studies: A Review of Regulatory Guidelines and a Survey of Current Practices
巻, ページ, 発行年	Toxicol Pathol.2007; 35: 742-750
分類 (①～④)	④
参照箇所の該当ページ	742-743, 749, 750

Society of Toxicologic Pathology (STP) は、各分野の規制ガイドラインの間には一定の違いがあることを認識しており、その中で、臓器重量の現行の評価方法が大きく異なることを認識している。そのため、STPは毒性試験における臓器重量に関する評価の方法を調査するためのワーキンググループを招集した。調査は、ヨーロッパ、北米、日本の製薬、獣医、化学、食品/栄養、消費者製品の種々の分野の企業にアンケートを配布し、意見を求めた。

調査では、げっ歯類及び非げっ歯類を使った様々な種類の毒性試験で日常的に測定されている臓器を特定し、臓器重量の評価法を比較し、臓器重量測定の有用性についての見解をまとめた。結果として、脾臓及び胸腺重量は免疫毒性、ストレス及び生理学的なかく乱の鋭敏な指標であると考えられ、さらに脾臓及び胸腺の病理組織学的変化は臓器重量の変化と相関があるとされた。一方、これらの重量測定にはほとんど価値がないという見解もあり、その理由として、胸腺では解剖時の技術的な問題や加齢に伴う退縮、ストレスに関連した影響などによるばらつきの大きさが挙げられた。脾臓では、動物間の変動性、ストレス関連の影響及び非げっ歯類 (特にイヌ) における安楽死に関連した脾臓のうっ血のような投与とは無関係の生理的な要因のため、その価値は限定的であると考えられた。さらにその価値が限定的であると考え人には、脾臓及び胸腺の重量変化はしばしば病理組織学的所見と相関しないことが多く、病理組織学的検査は被験物質投与による影響のより鋭敏な指標と考えられていた。

今回の STP 調査の回答では、一般的には、臓器重量は毒性試験における被験物質投与との関連性を評価するために有用なスクリーニングツールであると考えられていた。しかし、個別の臓器ごとの評価に関しては、実施した試験の目的に応じてその有用性の重要度については意見が大きく異なっていた。また、臓器重量データは、体重変化、被験物質の薬理作用、臨床病理学的データ、動物の絶食状態または絶食した場合の情報及び肉眼的・組織学的所見などのデータを含めた試験全体の流れの中で評価すべきであると強調する意見もあった。

番号	4-14
著者名	Marino DJ (2012b)
文献名	Age-Specific Absolute and Relative Organ Weight Distributions for Fischer 344 Rats
巻, ページ, 発行年	J Toxicol Environ Health A. 2012; 75(24): 1484-1516
分類 (①~④)	④
参照箇所の該当ページ	Abstract

フィッシャー344 (F344) ラットは、米国国立がん研究所 (NCI) 及び米国国家毒性プログラム (NTP) による毒性試験に使用される標準的なラット系統である。臓器重量の影響は、規制当局がヒトの健康リスク評価の毒性参照値 (TRV) を設定するために使用する重要なエンドポイントであり、加えて相対臓器重量を利用する生理学的薬物動態 (PBPK) モデルが TRV の策定に使用されるようになってきている。このように F344 ラットの集計された臓器重量は TRV 設定と PBPK モデリングに有益なデータになり得るため、本研究では、NCI/NTP の混餌、飲水、吸入試験から、被験物質非投与の対照群の F344 ラットの利用可能なすべての絶対及び相対臓器重量データを収集し、週齢別にまとめた。

得られた知見が示しているのは、月齢に応じて絶対重量が増加し、相対重量が減少するのは脳、肝臓、右腎、肺、心臓、甲状腺、右精巣である。結果によると、絶対臓器重量における一般的な変動傾向は、小さいほうから脳、右精巣、心臓、右腎、肝、肺、胸腺、甲状腺の順であることが示唆される。

番号	4-15
著者名	Marino DJ (2012a)
文献名	Age-Specific Absolute and Relative Organ Weight Distributions for B6C3F1 Mice
巻, ページ, 発行年	J Toxicol Environ Health A. 2012; 75(2): 76-99
分類 (①～④)	④
参照箇所の該当ページ	Abstract

概要

B6C3F1 マウスは、米国国立がん研究所 (NCI) 及び米国国家毒性プログラム (NTP) が実施する毒性試験に使用される標準的なマウス系統である。臓器重量データは、規制当局がヒトの健康リスク評価の毒性参照値 (TRV) を設定するために使用する重要なエンドポイントであり、加えて相対臓器重量を利用する生理学的薬物動態 (PBPK) モデルが TRV の設定に使用されるようになってきている。このように B6C3F1 マウスの集計された臓器重量は TRV 設定と PBPK モデリングに有益なデータになり得るため、本研究では、NCI/NTP で実施した試験から、被験物質非投与の対照群の B6C3F1 マウスの利用可能なすべての絶対及び相対臓器重量データを収集し、週齢別にまとめた。

結果によると、月齢に応じて絶対重量が増加するのは、肝、右腎、肺、心臓であり、相対重量が安定するのは、脳と右精巣であった。結果から示唆されるのは、絶対重量における一般的な変動傾向は、小さいほうから脳、右精巣、右腎、心臓、肝臓、肺、脾臓、胸腺であった。

番号	4-16
著者名	Dorso L, et al
文献名	Variability in Weight and Histological Appearance of the Prostate of Beagle Dogs Used in Toxicology Studies
巻, ページ, 発行年	Toxicol Pathol.2008; 36: 917-25
分類 (①~④)	④
参照箇所の該当ページ	Abstract, 923, Figure 13

概要

2つのブリーダーの生後23週齢から108週齢までの未処理の雄のビーグル犬の前立腺の重量と組織学的所見を評価した。27のガイドライン試験で得られた125頭の対照ビーグル犬のデータを使用した。年齢, 最終体重及び前立腺重量を測定した。前立腺について病理組織学的変化(腺房の発達, 分泌量, 拡張や炎症のパターンなど)を観察し, 必要に応じて等級付けを行った。前立腺の重量及び組織学的所見に及ぼす年齢, 最終体重及びブリーダーの影響, 及び動物間の変動の程度を評価した。

前立腺は精巣から分泌されるテストステロンによって発達することから, 精巣及び精巣上体の組織学的検査により性成熟の評価を行った(図1)。その結果, 前立腺の成熟度と精巣及び精巣上体の成熟度との間には強い相関関係があることが明らかとなった。さらに, Harlan Franceの31-40週齢の動物は90%が性成熟していたのに対し, Marshall Farms USAの同じ週齢の動物では10%しか性成熟しておらず, ブリーダーにより性成熟期のずれがあることが示唆された。同じ週齢の平均最終体重はMarshall Farms USAよりHarlan Franceの動物の方が統計学的に有意に高く, これは生産コロニーの遺伝的背景が異なるためと考えられた。これらのことから, 前立腺重量のばらつきの要因をまとめると以下のようになる。

- ① 性成熟期の個体差 (Marshall Farms USAの動物は41-50週齢, Harlan Franceの動物では31-40週齢の間に性成熟に達した)
- ② ブリーダーの違い (Harlan Franceの動物はMarshall Farms USAより同じ週齢でも体重が重く, 早く性成熟を迎えていたが, 同じ成熟段階の前立腺にブリーダー間の組織学的な差はなかった)
- ③ 同じ性成熟段階における個体差

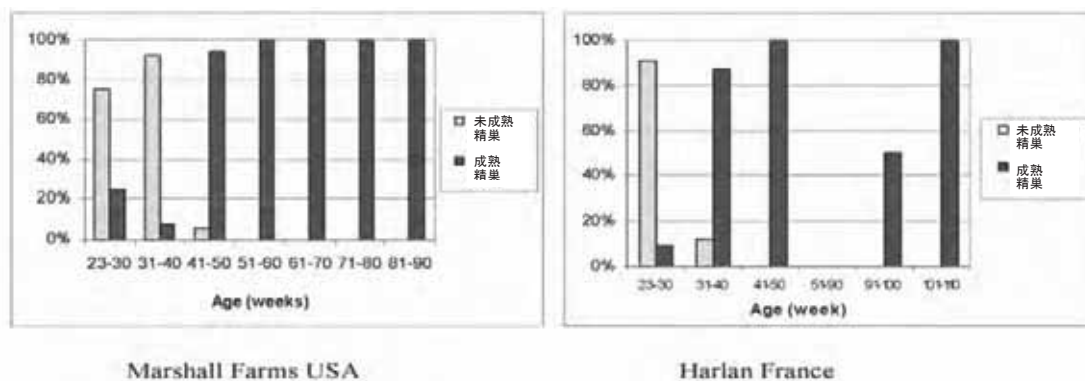


図1. 週齢別の未成熟及び成熟精巣の割合: Harlan Franceのイヌでは, 未成熟と診断された最高齢のイヌは31週齢であったのに対し, Marshall Farms USAのイヌでは41週齢であった。この結果から, Harlan FranceのイヌはMarshall Farms USAのイヌよりも早く性成熟を迎えることが確認された (Marshall Farms USA [n=97], Harlan France Harlan France [n=28])。

番号	4-17
著者名	Bailey SA, et al
文献名	Relationships Between Organ Weight and Body/Brain Weight in the Rat: What Is the Best Analytical Endpoint?
巻, ページ, 発行年	Toxicol Pathol. 2004; 32: 448-466
分類 (①~④)	④
参照箇所の該当ページ	Abstract, 461, 466, Table 3

概要

本研究では 1992 年 1 月から 1999 年 7 月までの期間に Wyeth Research で実施された, CrI:CD^B[SD]BR (Charles River Laboratories) ラットを用いた 26 本の 1 ヶ月間試験の対照動物 (雄 307 匹, 雌 266 匹) において, 様々な臓器がどの程度, 臓器重量と体重あるいは脳重量との間に比例関係があるかを調査し, これらの結果に基づき, どのエンドポイント (臓器絶対重量, 臓器重量/体重比または臓器重量/脳重量比) が標的臓器毒性を正確に検出できるかを検証した。

その結果を表 1 に示す。雌雄の肝臓及び甲状腺重量は体重との間に比例関係が存在し, 雌雄の副腎及び雌の卵巣重量では脳重量との間に比例関係が認められた。従って, 臓器重量/体重比は肝臓及び甲状腺重量を評価するのに適しており, 臓器重量/脳重量比は副腎及び卵巣重量を評価するのに適していると考えられた。脳, 心臓, 腎臓, 下垂体及び精巣の臓器重量については絶対重量, 臓器重量/体重比または臓器重量/脳重量比を用いた解析に適合していないことから, 共分散分析のような別の統計手法を利用すべきであろう。

表 1. 絶対臓器重量, 臓器重量/体重比及び臓器重量/脳重量比の最適な利用

臓器	体重比	脳重量比	別の統計手法
副腎		✓	
心臓			✓
腎臓			✓
肝臓	✓		
下垂体			✓
甲状腺	✓		
精巣			✓
卵巣		✓	
脳			✓

✓統計解析で, このエンドポイントに相当するモデルがデータに適合していることを示す。

このエンドポイントは, 群間に体重差が存在する場合に臓器重量差の評価に推奨される。

✓統計解析で, モデルが絶対重量, 臓器重量/体重比または臓器重量/脳重量比の解析に適合していないことを示す。臓器重量の解析には, 別の統計手法を利用すべきである。

番号	4-18
著者名	Sellers RS, et al
文献名	Society of Toxicologic Pathology Position Paper: Organ Weight Recommendations for Toxicology Studies
巻, ページ, 発行年	Toxicol Pathol. 2007; 35: 751-755
分類 (①~④)	④
参照箇所の該当ページ	752, Appendix 1

毒性試験における臓器重量の評価は、医薬品、化学物質、医療機器の評価に不可欠である。米国毒性病理学会は、7日間から1年間のGLPでの一般毒性試験において重量測定が推奨される臓器を提案する(付表1)。

脾臓及び胸腺はげっ歯類を対象とした試験では測定すべきである。胸腺重量は非げっ歯類でも価値があるが、生理的な胸腺の退縮は思春期以降のイヌやサル、特に期間が3ヶ月を超える試験では解釈を複雑にする。さらに非げっ歯類の脾臓重量は麻酔の方法や放血の程度により影響を受けることがある。脾臓及び胸腺重量は、リンパ系臓器重量に固有のばらつきがあるため、常に病理組織学的結果と併せて解釈すべきである。

付表1. GLPの一般毒性試験において重量測定が推奨される臓器 (投与期間は7日間から1年間*)

ラット	マウス	非げっ歯類
肝臓	肝臓	肝臓
腎臓	腎臓	腎臓
心臓	心臓	心臓
副腎	副腎	副腎
脳	脳	脳
精巣	精巣	精巣
前立腺		
精巣上部		
脾臓	脾臓	
胸腺	胸腺	
甲状腺/上皮小体		甲状腺/上皮小体
下垂体		下垂体

* : 発がん性試験は除く

番号	4-19
著者名	OECD (2001)
文献名	Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies
巻, ページ, 発行年	2002
分類 (①～④)	④
参照箇所の該当ページ	30-31, 33-34

体重変化, 食物及び水の消費量

試験期間中の個々の動物や動物群の体重変化（増加または減少）を、対照群の変化と比較することは、かなり重要な基準となる。体重変化は通常、飼料摂取量に関連しており、両者の評価を併せて行うことが必要である。体重の減少は必ずしもそれ自体が毒性と関連しているとは限らない理由としては、被験物質を含む飼料に対して動物の飼料に対する嗜好性が低下することにより、飼料摂取量に影響することがある。動物の忌避がみられた場合に餌の容器からの飼料のこぼしが増加することがあり、被験物質摂取量について配慮が必要となる。被験物質を飲料水によって投与する場合も同様に考慮する必要がある。

この忌避の作用は、多くの場合、最初の2～3週間の投与で明らかになるが、それ以降、これら動物は被験物質を含む飼料に適応できるようになり、体重増加が徐々に回復することがある。その際、亜急性毒性試験では、摂取した飼料1グラムあたりの個々の動物の増加量（食餌効率）が好転し、それに遅れて投与群の体重増加が続く場合があり、被験物質の毒性とは無関係に被験物質投与群と対照群間に投与初期に統計的に有意差を生じることがある。

被験物質の添加によって餌中の1つまたは複数の必須栄養素と相互作用することで、体重増加の減少や毒性反応の変化が生じることもある。この現象が亜急性毒性試験で観察された場合には、それに続く慢性毒性試験を開始する前に許容可能な手段（acceptable means）で克服することができる。まれに対照群の体重増加が低いために、他の群の増加値が高いようにみえることがあります。

飼料組成、飼料及び水の摂取量、体重の増加は、代謝、ホルモン、恒常性維持機構の変化や病歴、成熟などの動物の状態に重要な影響を与えることがあり、臓器やその他の重要なシステムに異常な所見がないにもかかわらず、これらの悪影響が観察された場合には考慮すべきである。

臓器重量の絶対値と相対値

これまで、亜急性毒性試験では、肝臓、腎臓、体重の変化が毒性のエンドポイントであるとされてきた。一般的には、臓器や体重の変化よりも病理組織学の方が最小中毒量を確定するのに適していると考えられているものの、臓器重量も通常、絶対臓器重量及び相対臓器重量（体重量ないし脳重量に対する相対）が評価に用いられる。

実験者が制御可能なものと制御できない両方の要因（概日リズム、食物摂取量、脱水、食餌の性質、動物の年齢、臓器の作業量、ストレス、殺処分方法など）も、臓器重量や体重に影響する。

殺処分時に放血処理をしなかった場合には、肝臓や腎臓の絶対あるいは相対重量を減少させるが、心臓、脳、脾臓重量には影響はないとされている。

被験物質が体重減少あるいは体重増加抑制に大きな影響を与えている毒性試験では、脳の相対量の

増加は評価上意味がなく、また、脳だけでなく他の臓器においても体重に依存して相対重量に影響した可能性があるため、これらは毒性影響による変化と区別すべきである。

また臓器重量変化に対する判断は、対照値との間の統計的に有意な差の有無だけでなく、病理組織学的データや代謝・トキシコキネティクスデータを含めて考慮する必要がある。