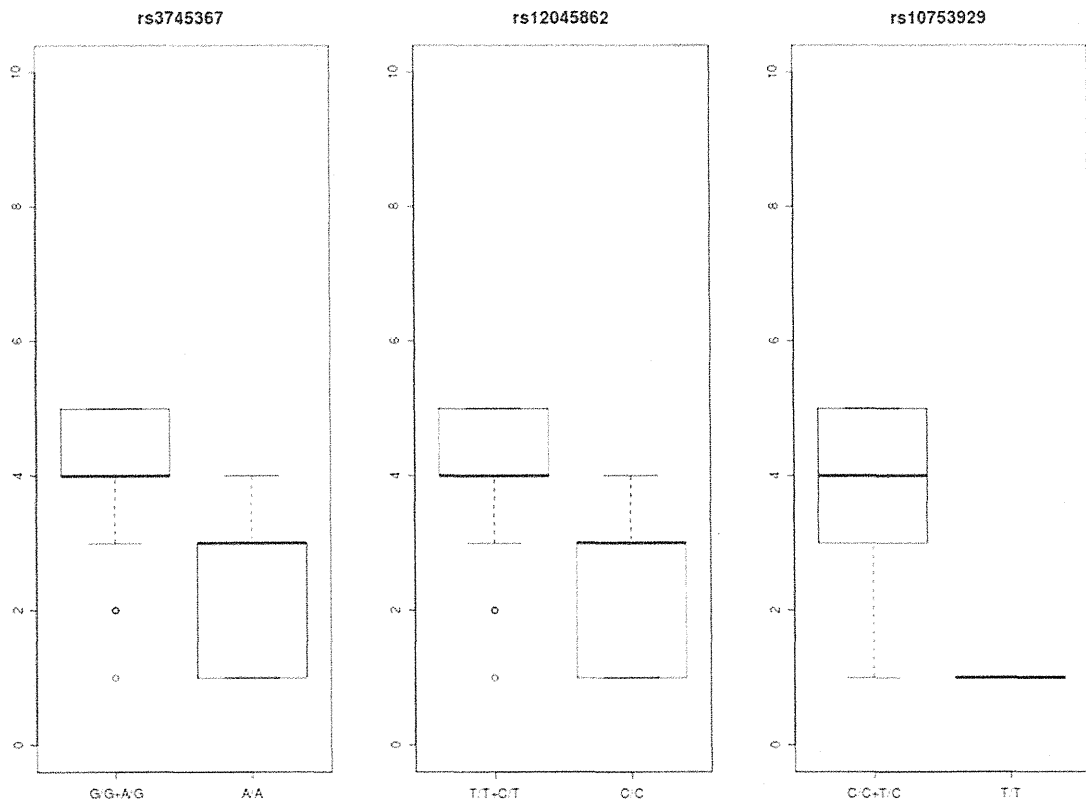


モデル 6-1 で有意となった SNP の遺伝継承と痛み強度の分布
劣性様式

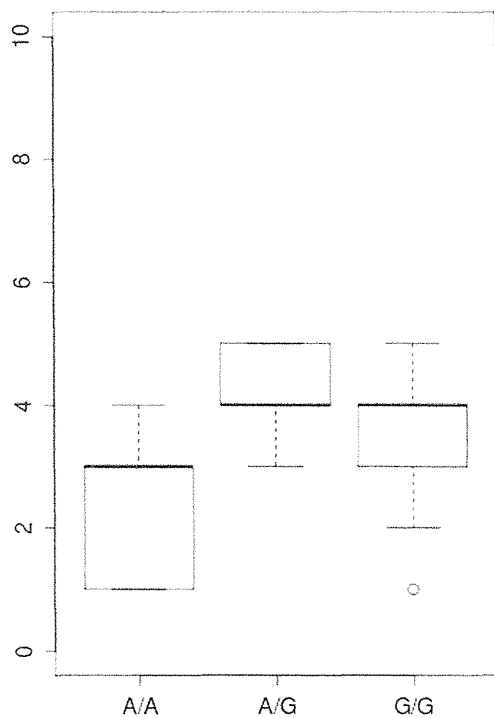
SNP		rs3745367	rs12045862	rs10753929
Gene		RETN	ADIPOR1	ADIPOR1
Allele	Major	G	T	C
	Minor	A	C	T
Phenotype	MajorHomo+Hetero	3.98±0.91 (N=57)	4±0.92 (N=59)	3.89±0.98 (N=64)
	MinorHomo	2.29±1.25 (N=8)	2.29±1.25 (N=7)	1±0 (N=2)
Mann-Whitney test		1.07.E-03	9.76.E-04	1.52.E-02



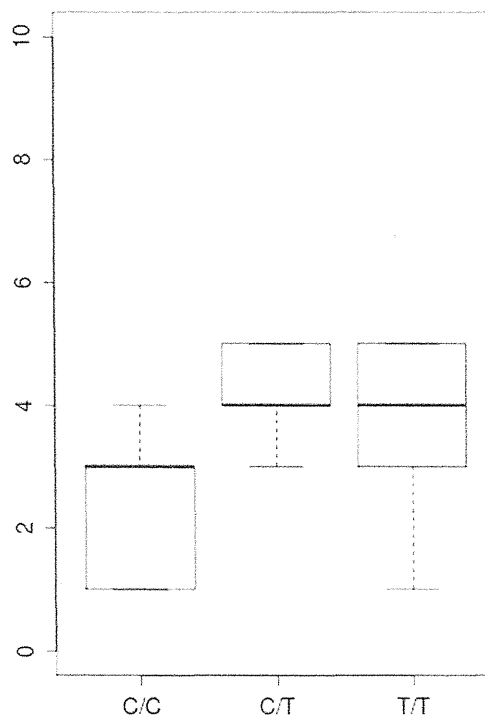
遺伝型様式

SNP		rs3745367	rs12045862
Gene		RETN	ADIPOR1
Allele	Major	G	T
	Minor	A	C
Phenotype	MajorHomo	3.79±0.98 (N=33)	3.84±1.04 (N=33)
	Hetero	4.24±0.77 (N=24)	4.25±0.64 (N=26)
	MinorHomo	2.29±1.25 (N=8)	2.29±1.25 (N=7)
Kruskal-wallis test		1.57.E-03	1.87.E-03

rs3745367

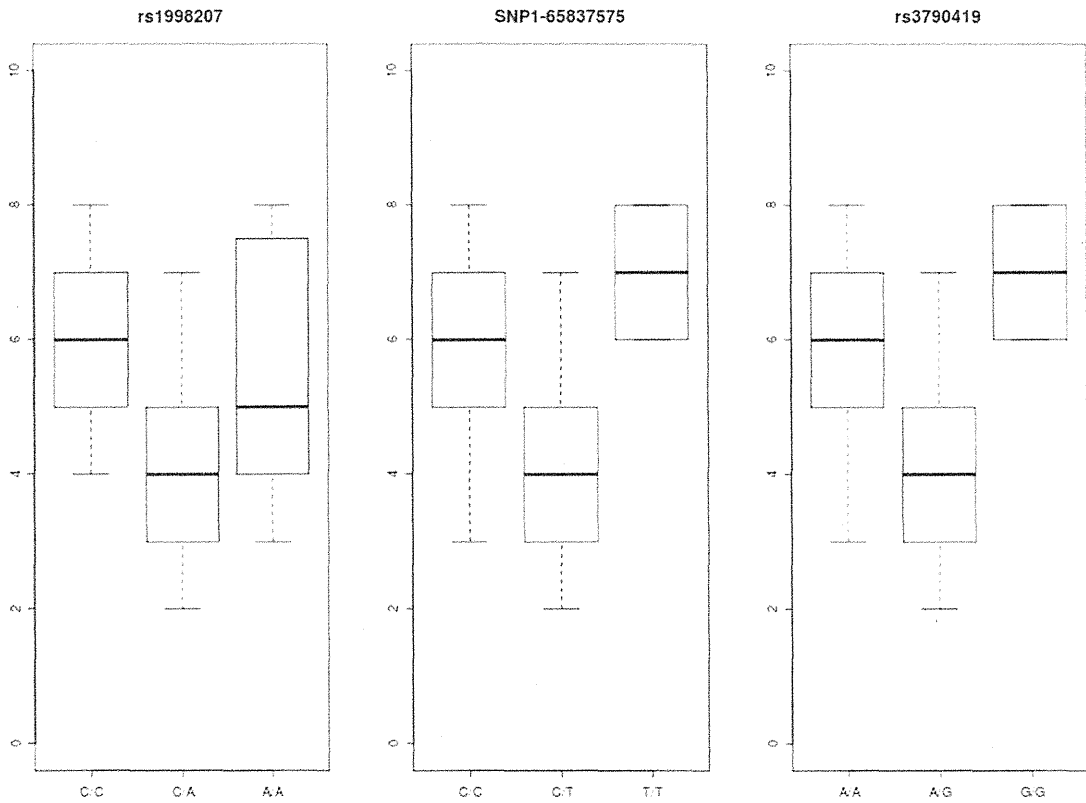


rs12045862



モデル 6-3 で有意となった SNP の遺伝継承と痛み強度の分布
遺伝型様式

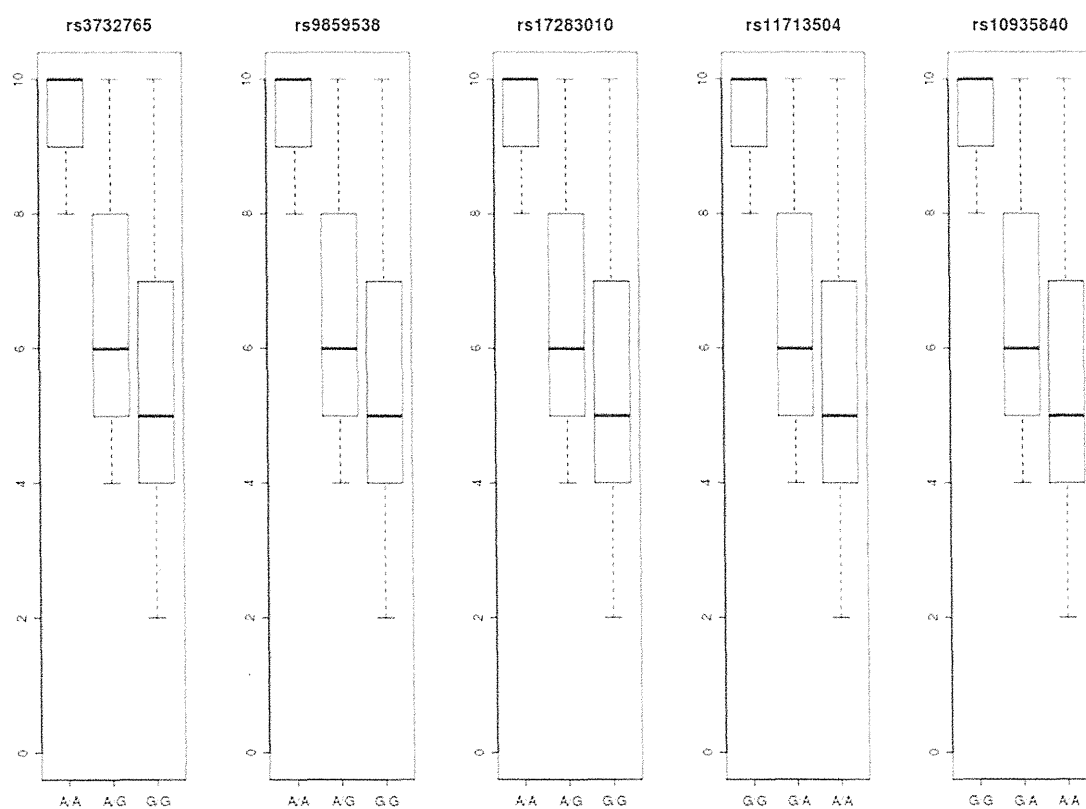
SNP		rs1998207	SNP1-65837575	rs3790419
Gene		SERPINA12	LEPR	LEPR
Allele	Major	A	T	G
	Minor	C	C	A
Phenotype	MajorHomo	5.5±2(N=24)	7±1.41(N=34)	7±1.41(N=34)
	Hetero	4.24±1.30(N=17)	4.08±1.44(N=13)	4.08±1.44(N=13)
	MinorHomo	6.16±1.24(N=8)	5.79±1.43(N=2)	5.79±1.43(N=2)
Kruskal-wallis test		7.36.E-04	2.49.E-03	2.49.E-03



モデル7で有意となったSNPの遺伝継承と痛み強度の分布

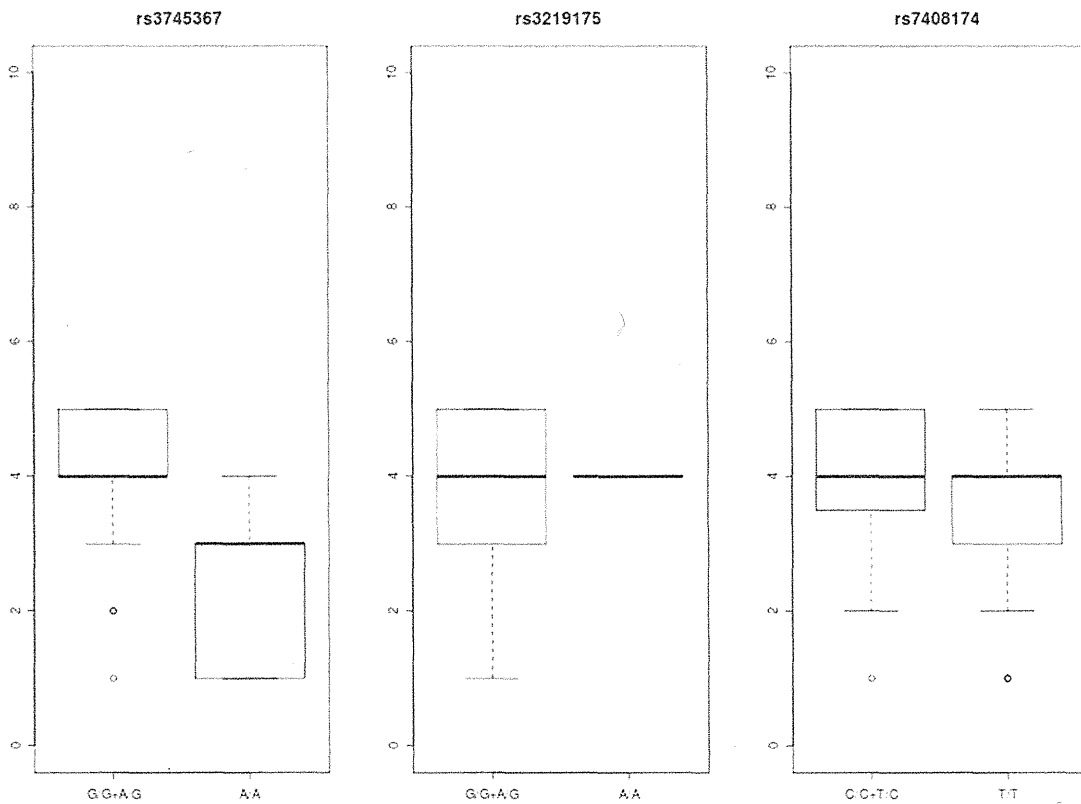
遺伝型様式、傾向性様式

SNP		rs3732765	rs9859538	rs17283010	rs11713504	rs10935840
Gene		P2RY12	P2RY12	P2RY12	P2RY12	P2RY12
Allele	Major	G	G	G	A	A
	Minor	A	A	A	G	G
Phenotype	MajorHomo	5.5±2(N=66)	5.5±2	5.5±2	5.5±2	5.5±2
	Hetero	6.4±1.57(N=21)	6.4±1.57	6.4±1.57	6.4±1.57	6.4±1.57
	MinorHomo	5.39±1.76(N=3)	5.39±1.76	5.39±1.76	5.39±1.76	5.39±1.76
Kruskal-wallis test		2.56.E-03	2.56.E-03	2.56.E-03	2.56.E-03	2.56.E-03
Wald test(傾向性検定)		1.12.E-04	1.12.E-04	1.12.E-04	1.12.E-04	1.12.E-04

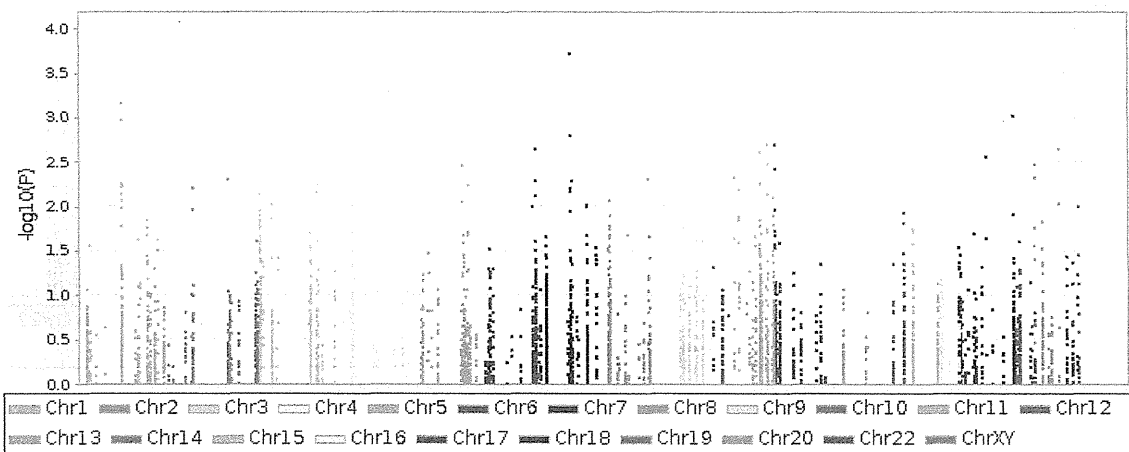
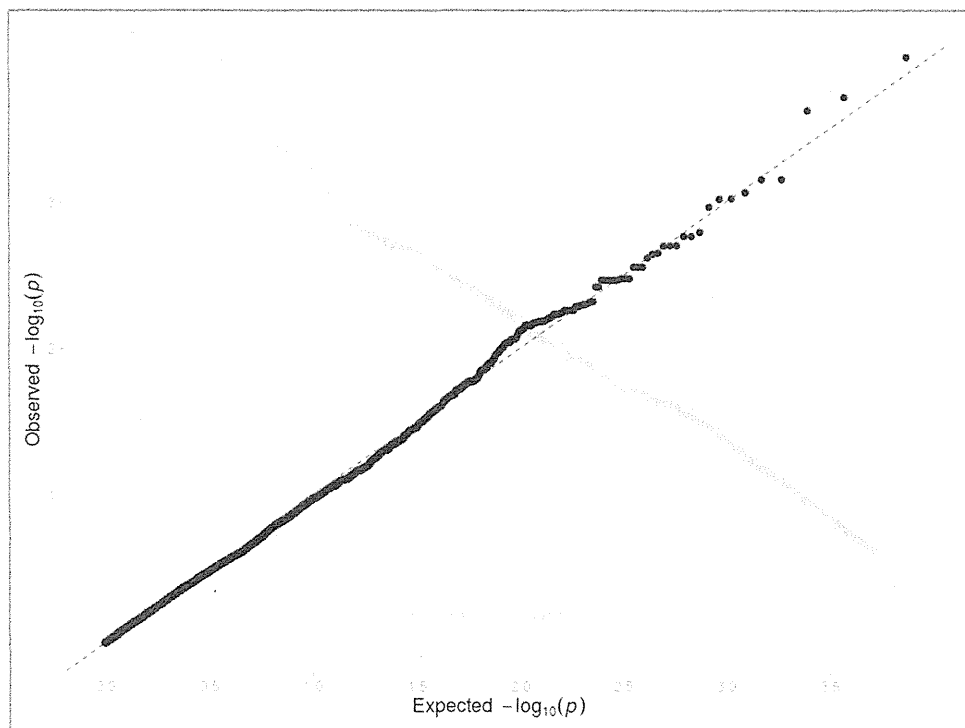


モデル 6-1 における RETN 遺伝子領域の分布 (劣性様式)

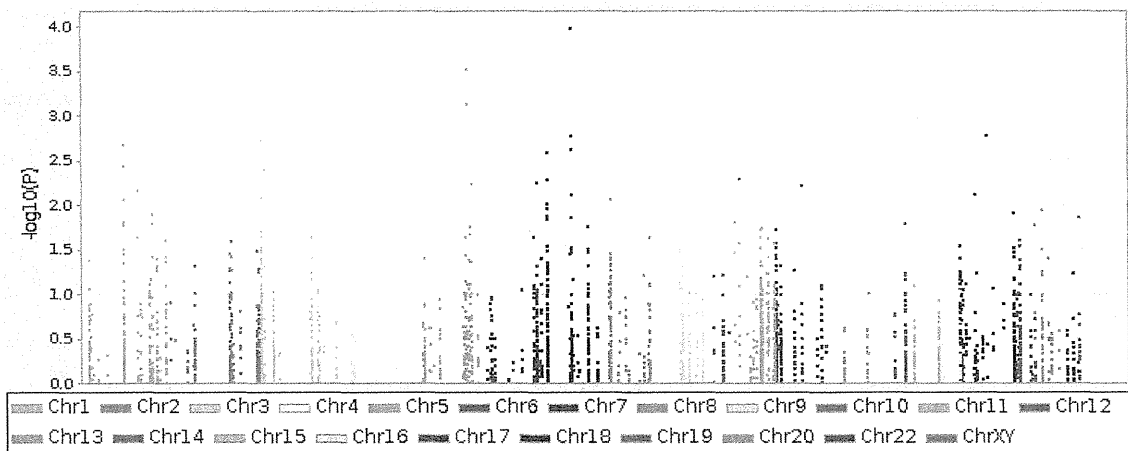
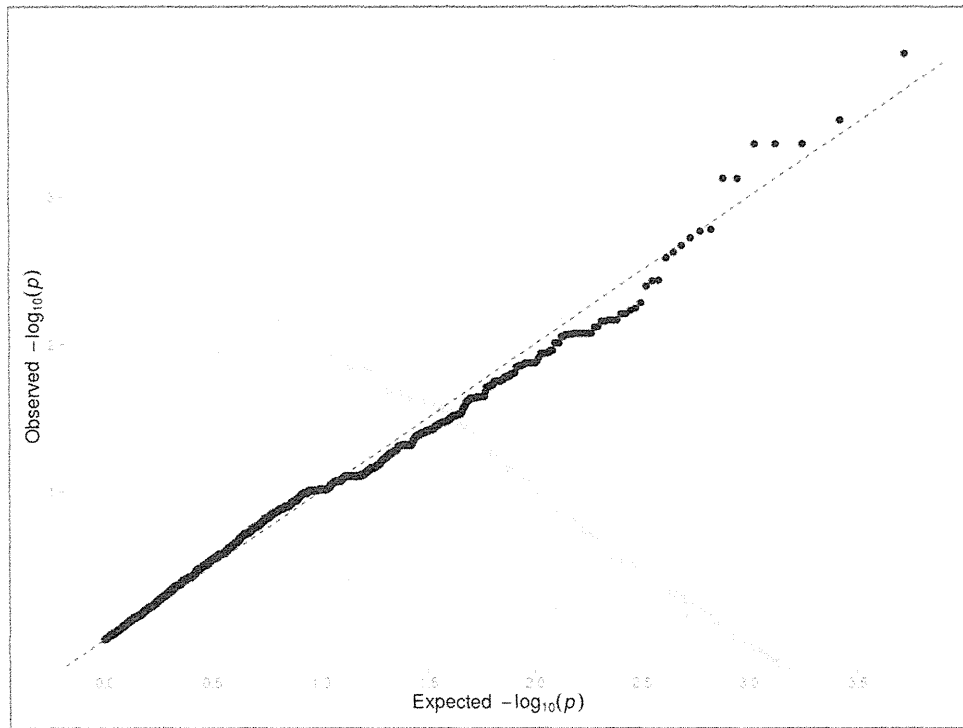
SNP		rs3745367	rs3219175	rs7408174
Gene		RETN	RETN	RETN
Allele	Major	G	G	A
	Minor	A	A	G
Phenotype	MajorHomo+Hetero	3.98±0.91(N=57)	3.79±1.11(N=65)	4±1.02(N=56)
	MinorHomo	2.29±1.25(N=8)	4±NA(N=1)	3.6±1.16(N=10)
Mann-Whitney test		1.07.E-03	9.50.E-01	1.63.E-01



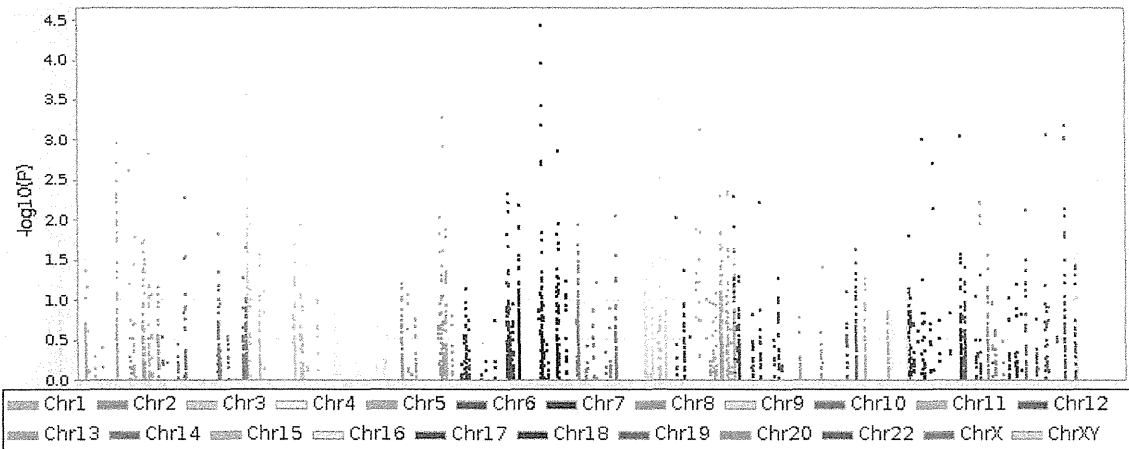
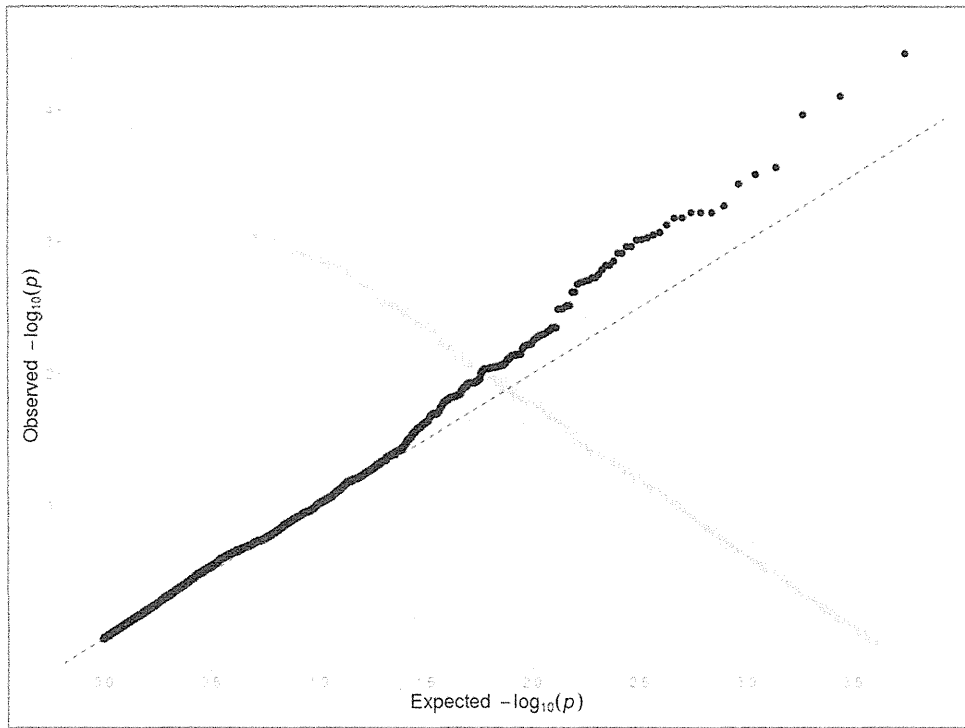
「疼痛強度」に対する候補遺伝子関連解析（「11221List of target genes.xls」ファイルの候補 SNPs1 ワークシートの遺伝子群）：侵害受容性疼痛である 49 個体
優性様式



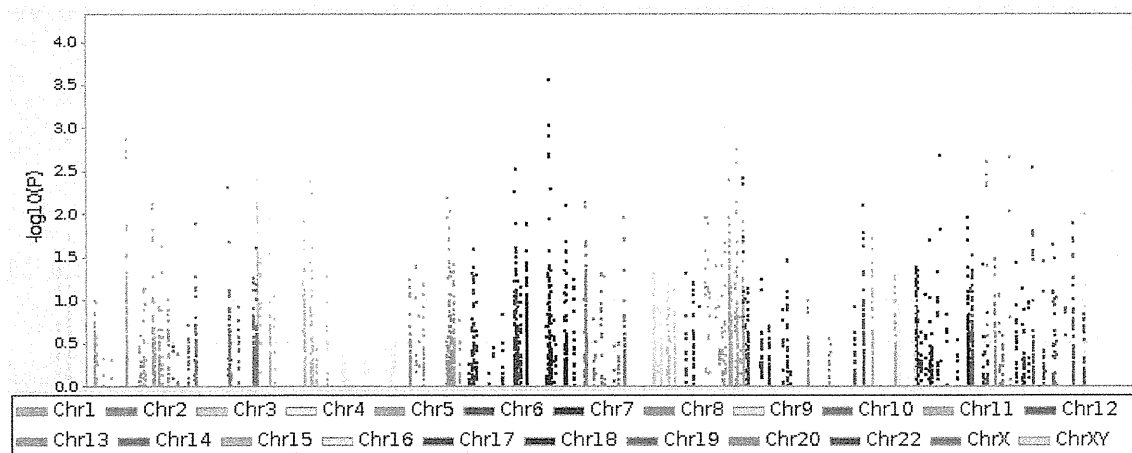
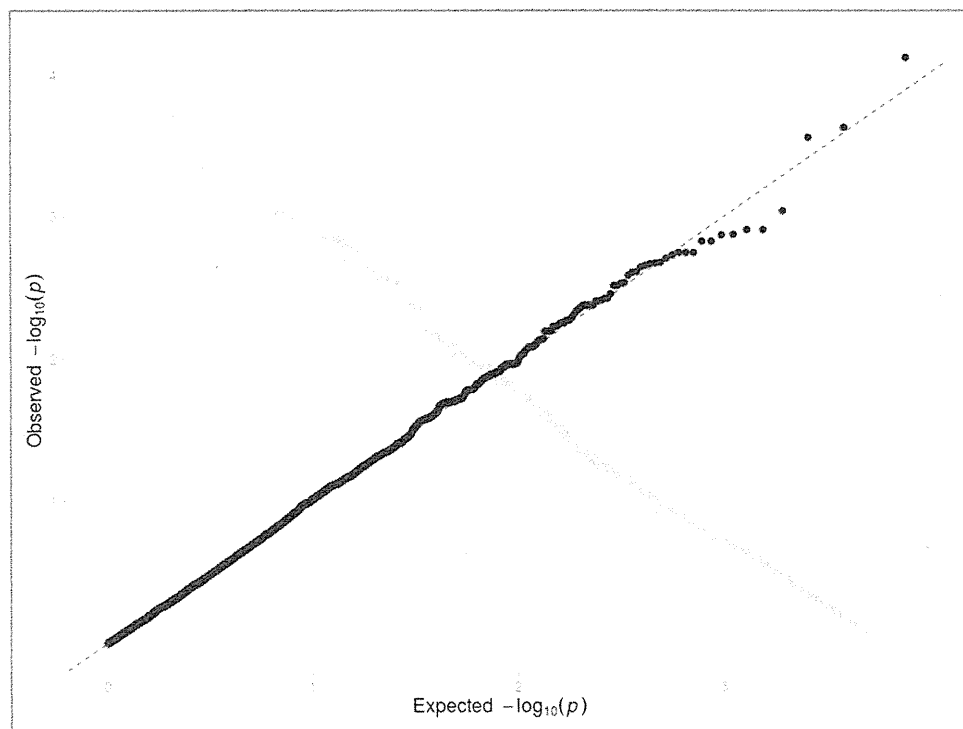
劣性様式



遺伝型様式

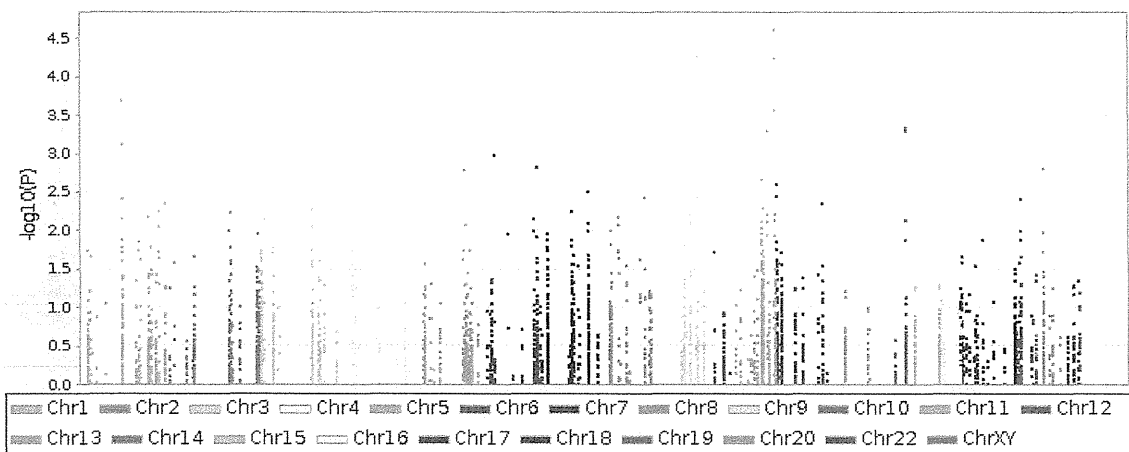
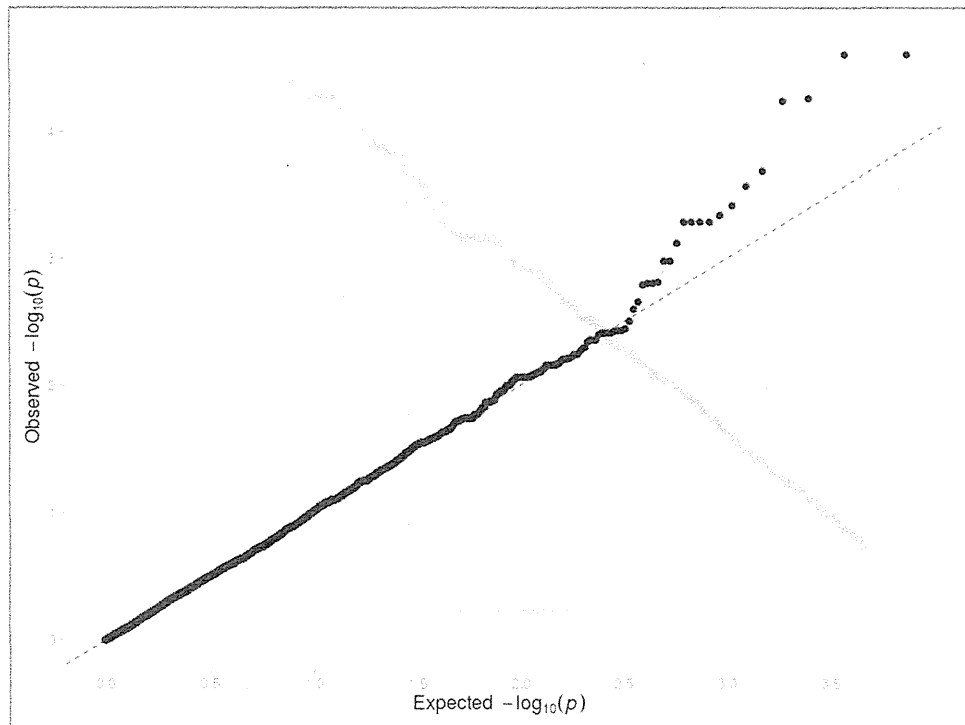


傾向性様式

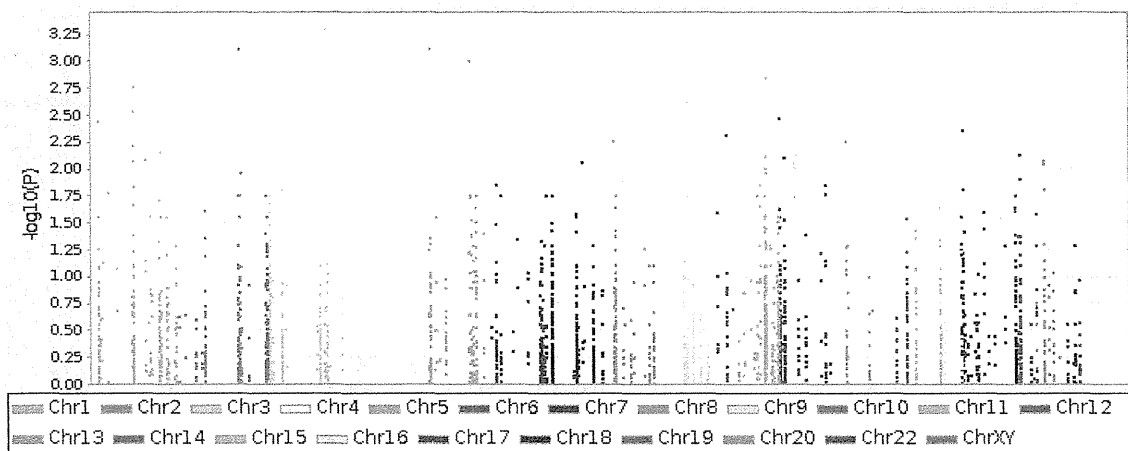
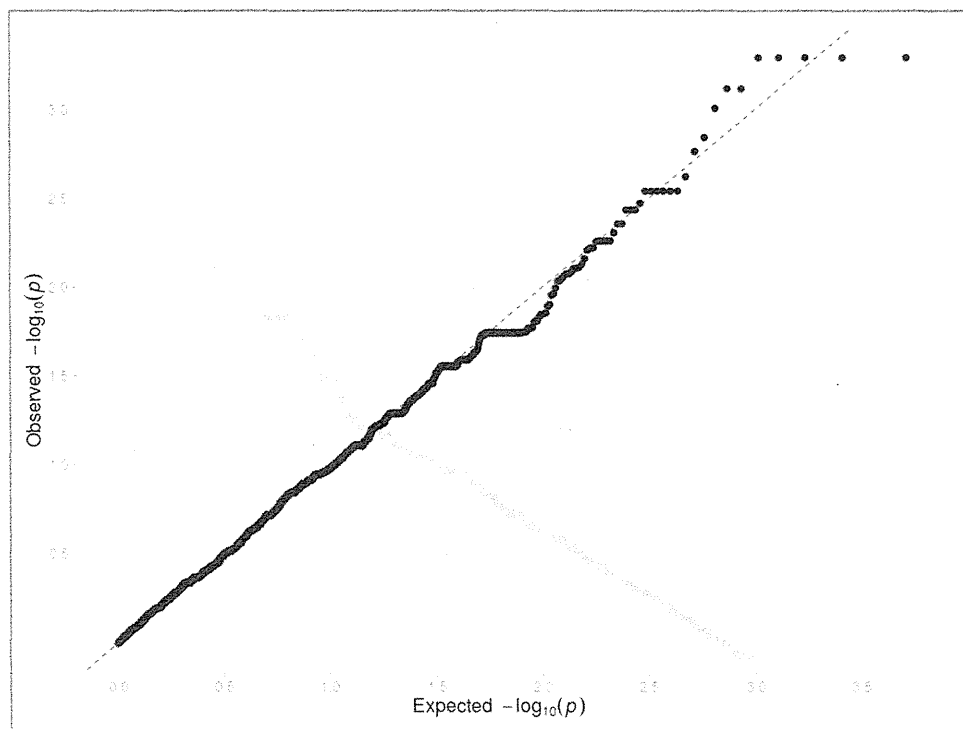


優性様式、劣性様式、遺伝型様式、傾向性様式において関連性を示している SNP はありませんでした。

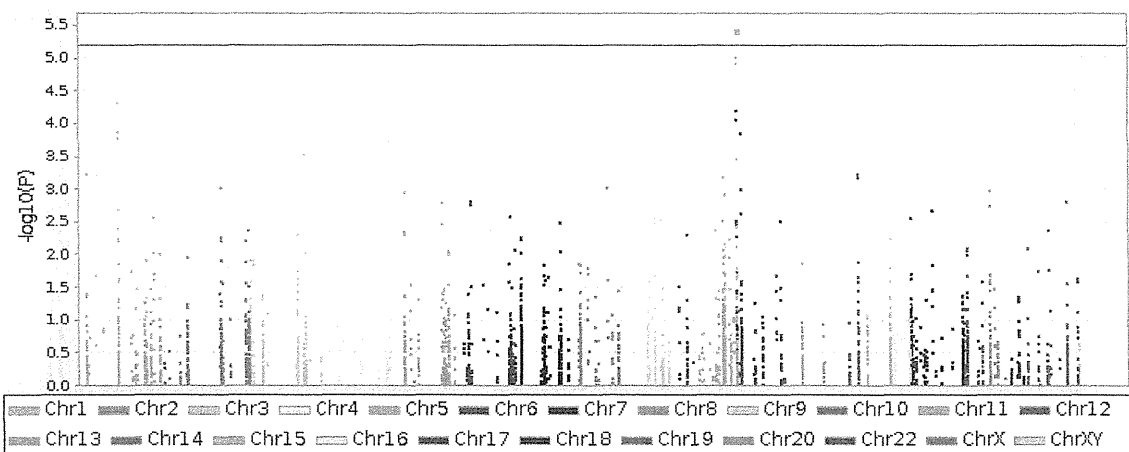
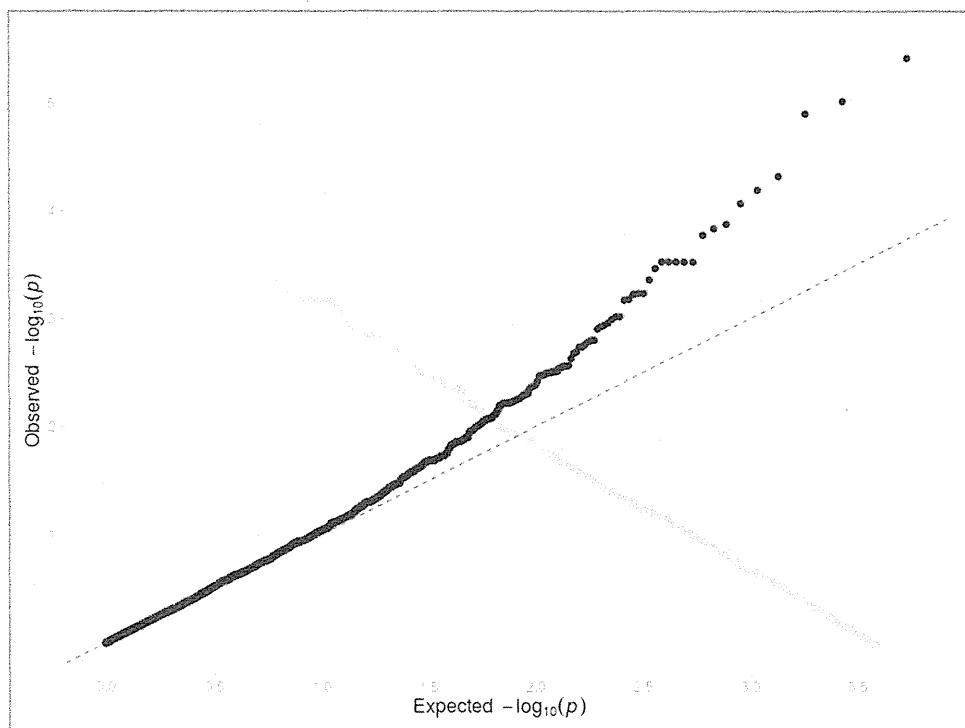
「疼痛強度」に対する候補遺伝子関連解析（「11221List of target_genes.xls」ファイルの候補 SNPs1 ワークシートの遺伝子群）：侵害受容性疼痛ではない 41 個体
優性様式



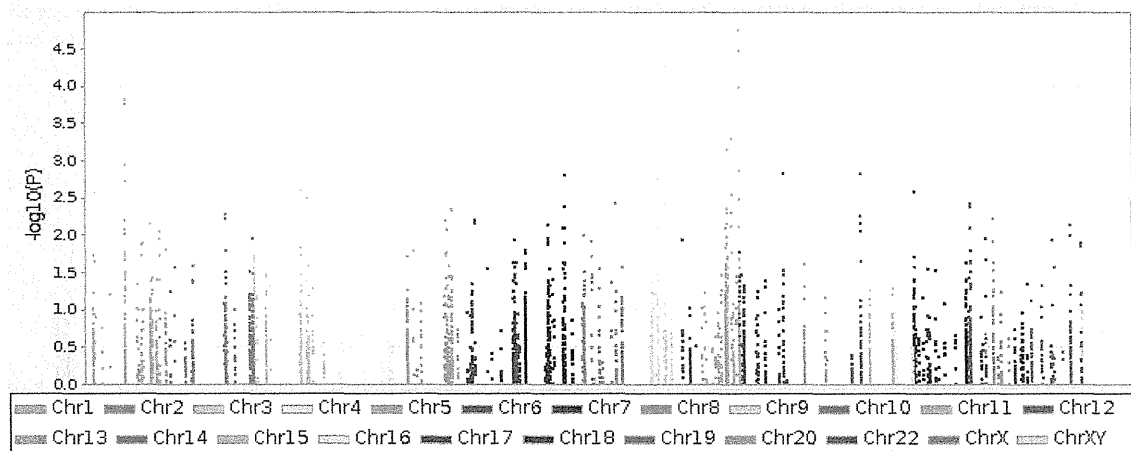
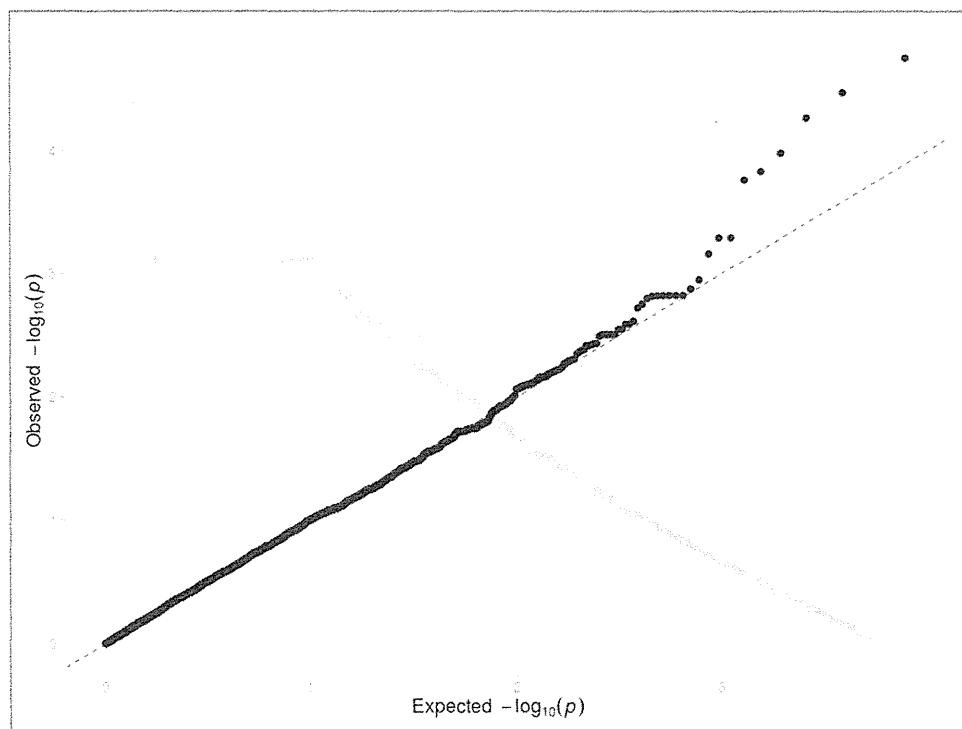
劣性様式



遺伝型様式



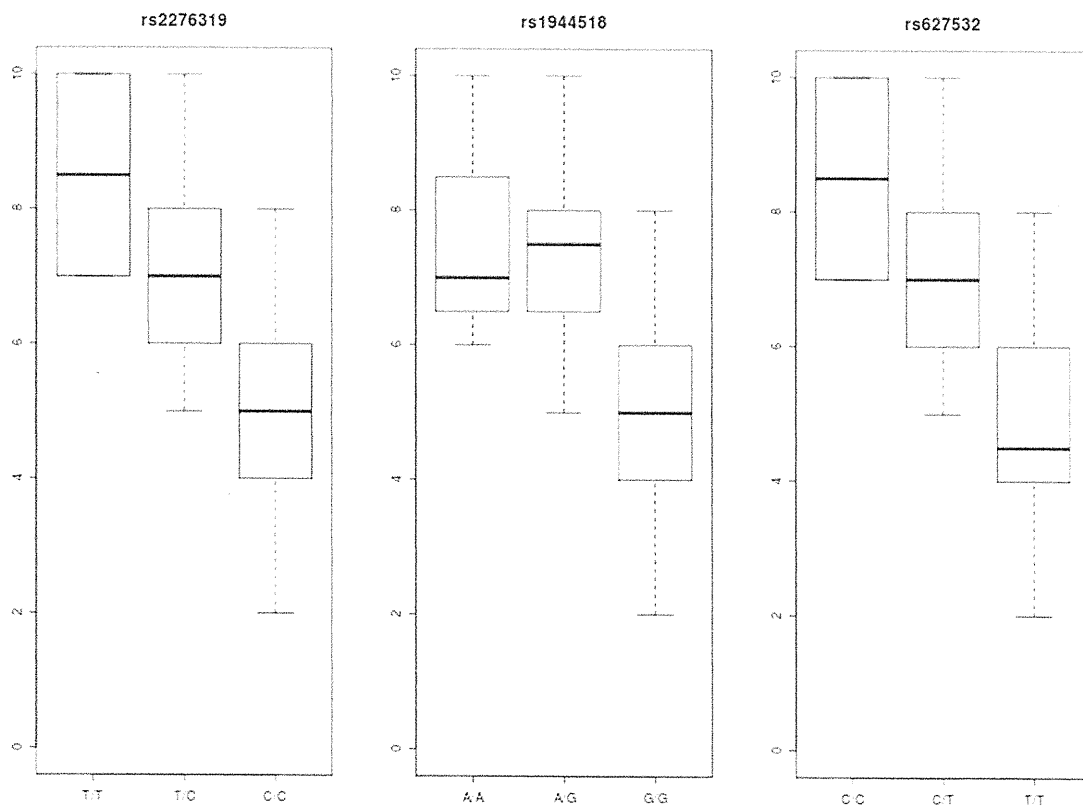
傾向性様式



優性様式、劣性様式、傾向性様式において関連性を示している SNP はありませんでした。遺伝型様式では BH 法で有意と判断された SNP が 3 つ、ボンフェローニの補正法によって有意と判断された SNP が 1 つ存在しました。有意となった SNP の情報は以下の通りです。

SNP	CHR	Gene	Location	Major Allele	Minor Allele	MAF	P	Bonf.Sig	BH.sig
rs2276319	11	GRIK4	120343031	C	T	0.26	3.85E-06	Yes	Yes
rs1944518	11	GRIK4	120348952	G	A	0.27	9.71E-06		Yes
rs627532	11	GRIK4	120331341	T	C	0.26	1.26E-05		Yes

SNP		rs2276319	rs1944518	rs627532
Gene		GRIK4	GRIK4	GRIK4
Allele	Major	C	G	T
	Minor	T	A	C
Phenotype	MajorHomo	5 ± 1.64(N=22)	5 ± 1.64(N=22)	5 ± 1.69(N=21)
	Hetero	7.41 ± 1.58(N=17)	7.5 ± 1.59(N=16)	7.28 ± 1.64(N=18)
	MinorHomo	8.5 ± 2.12(N=2)	7.67 ± 2.08(N=3)	8.5 ± 2.12(N=21)
Kruskal-wallis test		4.17.E-04	4.85.E-04	8.77.E-04



【Pain 2010; 149: 177-93】

Research design considerations for confirmatory chronic pain clinical trials: IMMPACT recommendations

Robert H. Dworkin^{*, et al.}

抄録

慢性疼痛への有効性が期待できることを示唆する過去の研究があるにもかかわらず、主要有効性解析で評価対象の治療とプラセボとの間で有意差がないことを示す慢性疼痛に関する臨床試験の数が増えている。これらの所見は、有効性の真の欠如あるいは有効性の証明に支障をきたす方法論的な面なども反映している可能性がある。慢性疼痛に関する臨床試験の間で重大な研究デザインの検討事項にかなりのばらつきがあることから、方法論的な弱点を明らかにし対処することによって新たな介入の鎮痛作用を裏証する可能性が強まるであろう。このような背景から、慢性疼痛に関する確認試験の研究デザインの重要な検討事項を特定し、これらの試験実施に関する勧告を行うことを目的としたIMMPACT コンセンサス会議が召集された。本稿では、試験の種類に加え、参加者の選択、試験相および試験期間、治療群および投与レジメンなどの慢性疼痛に関する確認臨床試験の主要要素に対する推奨事項を提示する。慢性疼痛に関する確認臨床試験の方法論的な面に注目し、研究を増やすことにより分析感度が高まり、最終的に慢性疼痛治療に関する有意義な評価を提供できると思われる。

キーワード: Chronic pain (慢性疼痛)、Randomized clinical trials (無作為化臨床試験)、Research design (研究デザイン)、Phase 3 trials (第 III 相試験)、Subject selection (被験者の選択)

1. 緒言

慢性疼痛治療に関する臨床試験の実施および解釈を促すために、Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) は、成人[29,143]、小児、青年[95]に関する一連のコア・アウトカム領域と指標に加え、改良アウトカム指標の作成[144]、個々の患者の改善および群間差の臨床的重要性の判定[30,32]、ならびに複数アウトカムの分析[145]の手法を推奨している。これらの推奨事項は、慢性疼痛試験でのアウトカム指標の選択、分析、解釈を含んでいるが、同じように重要な慢性疼痛に関する臨床試験で使用する研究デザインについての検討項目はない。鎮痛薬に関する臨床試験の試験デザインに大きな影響力のあった論文がほぼ 20 年前に発表されているが[94]、その後、慢性疼痛に関する無作為化臨床試験(RCT)の研究デザインに大きな進展がみられる。

二重盲検プラセボ対照 RCT は依然として、有効な治療介入を特定するための広く受け入れられている唯一の研究デザインであるが、慢性疼痛などの複雑な病態の検討に際しては大きな限界がある[62]。このデザインは、慢性疼痛を軽減するために設計された新規の介入について規制当局に承認を申請する際のピボタル試験で典型的に用いられている。しかし、そのような慢性疼痛試験で用いられる研究デザインおよび方法が注目されるようになったのは、最近のことである[27,31,68-70,107,108]。これは、主として臨床試験の数が増加した結果であるが、先行する前臨床研究や臨床研究で有効性が期待されることが示唆されているにもかかわらず[14,16,103,132,133,140,148,153 など]、これらの臨床試験で評価対象薬剤は主要有効性解析においてプラセボとの間で有意差を認めなかった。

これらのネガティブな所見が一般的条件または特定条件の下で有効性が真に欠如していることを反映しているかどうか、他の因子(わずかな治療ベネフィットの検出力が不十分、プラセボ群の奏効率が高過ぎる、データが欠落し補完を必要とした脱落例が多い、など)が有効性の証明を難しくしたかどうか、あるいはその両方が関与していたかどうかは明らかでない。慢性疼痛に関するRCTの研究デザインにはかなりのばらつきがあるため、これらの試験の方法論的特徴が有効な治療のベネフィットの証明に関与していた可能性もある。これらの結果の一部が研究の失敗を反映しているとする仮定は、慢性疼痛についての 2 件以上の試験で有効性が証明されていた薬剤に関する最近のネガティブな RCT によって裏付けられている[153]。複雑な慢性疼痛の病態も、臨床試験から得られる情報の種類に限界があるようである[62]。重要なことは、真に有効な介入に関するネガティブな試験は、既存治療に難治性の患者に有用かもしれない治療の特定を妨げるだけでなく、科学的・金銭的なリソースもかなり消費される点である。

これらの理由から、研究デザインの改善が臨床的に有用な治療の特定能の強化につながることを期待して、慢性疼痛治療に関する RCT の研究デザインにおける重要な検討事項を特定し、これらの試験実施に関する勧告を行うことを目的とした IMMPACT のコンセンサス会議が召集された。本稿では、「確認(confirmatory)」または「第 III 相(Phase 3)」と称した試験、すなわち、最低でも有効性および安全性に対する予備的エビデンスが既に存在する治療の有効性および安全性について、確かなエビデンスを得るために実施された適切な対照群のある(well-controlled)試験に注目した。有効性についての最初のエビデンスが得られる臨床試験デザイン(すなわち、「概念実証[Proof-of-concept]」または第 II 相[Phase 2]試験)に対する推奨事項は今回の IMMPACT 会議のテーマであり、別に報告す

る予定である。

本稿では、慢性疼痛に関する確認臨床試験のための研究デザインの検討事項を示すことを目的とし、以下の主要項目を含む：(1) 研究被験者、(2) 試験相、(3) 治療群、(4) 確認試験の種類。慢性疼痛に関する臨床試験の分析および解釈に関連した種々の事項については、他に報告されている[30,32,145]。本稿の基となった IMMPACT 会議では薬物療法の確認試験に注目しているが、多くの検討事項は他の慢性疼痛治療（理学療法、脊髄刺激、心理社会的介入など）に関する試験や異なる治療法の併用に関する研究などの臨床試験も含んでいる[78]。頭痛に関する臨床試験の研究デザインは、本稿の範囲外である[131]。

2. コンセンサス会議の手順

IMMPACT コンセンサス会議は、大学、政府機関、産業界、患者擁護団体からの参加者による国際的グループが参加して開催された。参加者は、慢性疼痛治療に関する試験のデザインおよび評価に関する研究、臨床に上または行政上の専門性に基づいて選択された。率直で効率的な議論を推進するために会議の規模を制限し、種々の学問分野および専門性を幅広く代表する人々を含めることとした。議論およびコンセンサスに基づく勧告の開発を促すために、会議では補足的講義で以下が検討された：(1) 上乗せデザインに焦点を合わせたてんかんに関する臨床試験 (JAF)、(2) プラセボ群の反応に焦点を合わせた抑うつに関する臨床試験 (Michael E. Thase, MD)、(3) 臨床試験での効果の発現および持続性の分析および解釈 (SE)、(4) 強化した組入れによる無作為化治療中止試験 (NPK)。以下の推奨事項の多くは系統的な研究に基づいたものではなく、むしろ慢性疼痛やその他の疾患（抑うつ、てんかん、など）の治療に関する臨床試験の方法および結果からの外挿であることを強調することは重要である。

3. 研究対象

臨床試験の組入れ基準および除外基準は参加者の特徴を設定するものであることから、広範囲な集団に対する一般化に影響を及ぼすだけでなく、試験が有効性を首尾よく証明できる可能性にも重大な影響を及ぼす[115,116]。これらの 2 つの目標はしばしば対立する。例えば、地域社会で患者の不均一性を反映した状態の対象者を採用することにより試験の治療効果の検出力が低下する可能性がある。例えば、患者の人口統計学的特徴および疾患重症度、罹病期間、併存疾患、併用薬などの臨床的特徴に大きなばらつきがあれば、治療に対する患者の反応にも大きなばらつきが生じる。組入れ基準の設定では、臨床試験の「分析感度」を損なう可能性があるばらつきの既知の原因や潜在的原因を有する幅広い集団を反映した対象者の特徴に慎重なバランスを取ること、すなわち、試験がプラセボまたは実薬対照との比較で治療効果を証明する能力が要求される。

組入れ基準が極めて具体的な（例えば、HIV に関係した神経障害）確認試験で実薬治療の有効性を示した結果は、臨床診療の場でこの治療法が有効でないとと思われる患者にも（例えば、腰腿神経根障害）、この治療が行われる危険性がある。逆に、幅広い患者集団（例えば、あらゆる種類の神経障害性疼痛）を検討した試験での有効性欠如は、少なくとも一部の患者に有効であると思われる治療が放棄される危険性がある。この点においては、例えば、極めて特異的な病態に対する新たな化

合物の有効性を証明する「説明的」試験と、例えば、よくある不均一性の臨床症状（例えば、腰痛）に対する既知治療の有効性を評価する「実践的」試験の区別が役に立つ[127]。このように、臨床試験は予め慎重に設定された目的に基づいて治療を評価するように設計されるべきである（例えば、特定の適応に関する規制当局の承認とプライマリケアにおける有効性の推定を区別する必要がある）。

3.1. 人口統計学的特徴、その他の特徴

すべての臨床試験に関する出版物およびオンライン登録は、全被験者について少なくとも年齢、性別、人種/民族に関して報告する必要がある。これらの特徴の中で年齢のみが組入れ基準の 1 つとして頻繁に使用されている。例えば、18 歳未満の患者は初期の確認試験から除外されることが通例であり、帯状疱疹後神経痛 (PHN) のような小児や青年が罹患することがほとんどない特定疾患に関する臨床試験では除外できる。さらに、高齢患者 (85 歳超) の場合、併存疾患の有病率や薬物療法の施行率が高い人に安全性や忍容性の懸念があるならば、試験対象から除外されることがある。人口統計学的データを広範囲に報告することで、治療の有効性、安全性、忍容性が患者背景に応じて変化するかどうかについての研究が可能となり、そのような分析はどの患者がその治療に適応となるかを証明する重要な要素となる。

その他の社会経済的特徴、心理社会的特徴、生活の質に関する特徴も慢性疼痛患者の場合、特に重要となることがある。例えば、被験者の教育や職業の状況ならびに健康関連の生活の質は、薬物療法および非薬物療法に関する臨床試験において治療反応と関連していた[例 67,100]。さらに、疼痛に関連した現在訴訟中の患者や労働災害補償を受けている患者については、文書で証明されている疼痛の軽減が法的な状況または身体障害の状態に不利な影響を及ぼす可能性があり、それらの患者の疼痛改善の報告が妨げられることが懸念されるため、慢性疼痛に関する臨床試験から除外されることが多い。しかし、身体障害は慢性疼痛患者によくみられ、これらの患者を除外することは臨床診療の場での一般化に制限を加えることになるかもしれない。

治療アウトカムのばらつきの一因となる可能性がある既知の患者背景および潜在的な患者背景の徹底的な評価を臨床試験の開始に先立って実施すべきであり、これらは対象者を無作為割り付けする前に評価される必要がある[62]。一部の患者背景は疾患の進行や薬剤の代謝に影響を及ぼし、それによって治療反応にも影響を及ぼすことがある。例えば、体重は痛性糖尿病性末梢神経障害 (DPN) に影響を及ぼす可能性があり、腎機能低下は腎臓で代謝される薬剤の副作用を増加させる可能性がある。患者を代謝が遅い人と速い人に分類可能な場合もあり[134]、薬物反応の可能性が高い患者を特定するために有効性および安全性に関する分析でこれらの情報の検討が可能である。ヒトゲノムに関する知見が増えるにつれ、RCT で遺伝子型解析および有効性が検証されているバイオマーカーを使用することにより、薬剤と被験者を薬物動態学的および薬力学的な相互作用の非常に幅広い知識に基づいてマッチングさせることがさらに可能になると思われる。

3.2. 診断

正確な診断は、臨床試験に組入れられる被験者が治療対象疾患または状態を有しているかどうかを確認するために欠かせない。慢性疼痛症状はしばしば、このプロセスにおいて次の 2 つの要素が

ある：(1) 例えば、変形性関節症(OA)や糖尿病などの基礎疾患の診断、(2) 例えば、OA による膝関節痛や有痛性 DPN など疾患に起因する疼痛症状の診断。慢性疼痛を伴う疾患の多くは、広く受け入れられている診断基準がある。例えば、股関節の OA に対する米国リウマチ学会の基準では、股関節痛に加えて 2 項目以上の臨床検査による陽性結果および X 線上的所見が要求される[1,2]。

しかし、慢性疼痛症状の診断は時にきわめて難しい場合がある。最近改訂された神経障害性疼痛の定義は、2 種類の確認検査がある場合に限り「確定」診断が行われる評点方式がある[141]。神経障害性疼痛の原因となっている病態のいくつかに対して多様な診断アプローチが開発されているが[例 21,26,36,54,65]、確認検査ならびにその陽性および陰性予測の妥当性に関するエビデンスの入手は限られているようである。

このような妨害はあるが、病因的に異なるいくつかの疼痛症状については一般に広く受け入れられている診断基準が開発されており、多様な作用機序を有する治療の有効性を証明する際の基準となっている。これらの症状には、OA、関節リウマチ、DPN、PHN および線維筋痛症(FM)に関連した慢性疼痛がある。慢性腰痛は一般集団で非常に高頻度に見られるが、その診断基準はまちまちであり、神経障害性または非神経障害性の疼痛機序が治療の標的であるかどうかにかかわらず依存している。

特定の病因を標的とした治療に関する相次ぐ臨床開発は多くの場合、既存治療の有効性の証明に先立って作用機序や治療対象の根底にある病態が明らかにされている所見と一致している。そのため、早期の臨床試験が特定の慢性疼痛の症状に有効であるというエビデンスを示している場合、同一症状を対象として確認試験を実施することは妥当である。

診断的に均一性の高いサンプルでは有効性が証明できる可能性が高いと一般に考えられているが、これは不均一なサンプルは治療に反応しない症状のある対象者を含んでいる可能性が高いためである。従って、可能な限り診断手順および検査を確認試験で使用する前に標的集団においてそれらの精度を評価すべきであり、また明らかに標的症状を有する患者のみを組入れ、標的症状が存在する可能性しかない患者を除外できるような正確な診断プロセスであることが理想的である[122]。このことによって、試験結果が適用される集団の正確な特徴付けがさらに可能となる。

診断プロセスおよび関連手順を試験実施施設および試験責任医師の間で標準化することにより、より均一性の高い患者サンプルの確保が期待できる。診断が主として病歴や症状に基づく場合は、一貫性のあるアプローチを用いて医師は患者に問診を行い、標準化されたプロセスを用いてこの情報を捕捉すべきである。例えば、スクリーニングでの来院時に、症状を明らかにするために一連の問診票の使用が可能であるが、これらは全ての患者に対して同一の順番で示し、同一の文言を用いる必要がある。Patient-Reported Outcomes Measurement Information System (PROMIS)などの現在進行中のイニシアチブでは、患者の自己報告による疼痛および関連アウトカムを評価するために標準化されたアプローチを開発中であり、これによって、臨床試験との間で、また場合によっては異なる疼痛症状の間で評価の同等性が裏付けられると考えられる[18,110]。可能であれば、観察者間および観察者内の再現性および信頼性を前もって判定し、ばらつきの原因を少なくするための対策を講じるべきである。これらと同じような配慮を理学的診察および徴候の文書化の標準化でも適用すべきである。また、電気生理学的所見や X 線所見[例 2,90]などのさまざまな診断手順およびその結果の解釈についても、このような客観性が高いとされる手順であっても信頼できない可能性もあることから[56]、

これらの標準化にも適用すべきである。

病因に基づく診断が現時点では、慢性疼痛に関するRCTにおいて標準とされているが、慢性疼痛の診断、評価、治療に対する至適アプローチは、まず患者の疼痛が説明できる病態生理学的機序を明らかにし、これらの機序を至適治療アプローチ選択の根拠として使用することが提案されている[9,92,151,152]。しかし、慢性疼痛に関する確認試験の設計において依然として、これらの機序を十分に明らかにすることは出来ず、またどの治療がそれらの機序を標的とするかについても明確にできない。現在進行中のトランスレーショナル・リサーチ(translational research)は、慢性疼痛管理への機序に基づくアプローチの基礎となると考えられる。このことは、臨床試験に疼痛機序を特定するための系統的評価を含む必要があると思われる[例 10,112,113,126]。そのような機序に基づく治療へのアプローチが開発され、効果が検証されるまでは、現時点では根底にある機序(例 中枢性感作)よりも疾患に特異的な診断(例 膝関節 OA)を標的とした試験を続けている。

3.3. 罹病期間および疼痛持続期間

慢性疼痛試験では、研究対象者の慢性疼痛が自然に消失する可能性がある比較的軽いものであるかどうかを確認するために罹病期間の評価が必要である。慢性疼痛治療の有効性評価に際しては、疾患のばらつきに伴う変動を最小限に抑えるために基礎疾患の状態が安定している患者が一般に望ましい。しかし、特定の病態は重要な時間的変化を特徴としており(例 多発性硬化症)、これは試験デザインで対処しなければならない。典型的なものとして、慢性疼痛に関する確認 RCT の組入れ基準は、基礎疾患の診断からの最短期間および慢性疼痛の最短罹病期間を必要としている。例えば、有痛性 DPN に関する臨床試験では、対象者は DPN の診断を 1 年以上前に受けており、DPN に関連した疼痛を 6 ヶ月以上経験していることが必要とされる場合がある。

疾患の発症を知れば、慢性疼痛が特定状況の下で存在するかどうかを十分に検証できるかもしれない。例えば、帯状疱疹の発疹の発現日から、PHN の臨床試験で疼痛持続期間が簡単に判定できる[26]。しかし、他の多くの病態については、疾患持続期間は症状持続に関する患者の報告に基づいて確認しなければならない。疼痛持続期間に関する患者の報告は組入れ基準が満たされているかどうかを確認するために用いられることが多いが、患者によるこれらの報告の収集は必ずしも十分に開発された手法を用いて系統的に行われていない。診断コードまたは明らかな前治療歴(例えば、試験開始前の特定期間の鎮痛薬処方[125])を用いて症状持続期間に関する患者の報告を補強することにより、患者の回想の正当性を立証できるかもしれない。しかし、これらのデータは必ずしも入手できないことから、このような組入れ基準は選択バイアスの原因となるかもしれない。アプローチにかかわらず、臨床試験に関する出版物は期間判定に用いられた具体的な方法を明確に記述すべきである。

3 ヶ月以上持続している疼痛は慢性とみなされ[97]、多くの臨床試験でこの期間が組入れ基準の 1 つとされている。しかし、6 ヶ月以上の疼痛持続を必要とすることによって、自然経過の結果、疼痛の著しい軽減を経験する可能性がやや低い対象者サンプルが得られるであろう。このように厳しい疼痛持続期間を用いると組入れスピードを遅らせることにもなるため、既存のデータベースを検討して疼痛持続期間が比較的短い被験者を解析対象から除外した場合に分析感度が高まるかどうかを判定す

る必要がある。

個別のカットオフ・ポイントについては任意に設定されているが、5～10年の罹病期間または疼痛持続期間の上限[例 51,108]も慢性疼痛に関する一部の試験で用いられている。このような疼痛持続期間の上限を用いることで、前治療の効果がみられない難治性の患者や治療効果があるとは思えない非常に長い疼痛持続期間を有する患者の除外に有用であると思われる。しかしながら、治療が不十分であった患者や既存の療法と比較して異なる作用機序のある新たな介入に反応すると考えられる患者が除外される可能性もある。データが欠落している場合には、疼痛症状の病態生理学的機序、既存の治療と比較して異なる治療の作用機序、十分に確立された疼痛患者の多くの除外が組み入れに及ぼす影響を慎重に検討した上で、罹病期間または疼痛持続期間の上限を慢性疼痛試験に組み入れるべきである。

3.4. 疼痛強度

総合的または平均的な疼痛強度は、同一の疼痛症状を有する患者の間でもかなりのばらつきがある。このばらつきには、疼痛感受性に対する個人差[34,41]、活動レベル、対処能力、基礎疾患または疼痛原因の重症度など複数の原因がある。現在のところ、慢性疼痛に関するほぼ全ての試験で、ベースライン時の患者の自己報告による疼痛強度が一定以上(多くの場合、0～10の数値化スケール[NRS]で4)であることを要求している[7,90,118]。一般に、このような測定尺度の心理測定的特徴は、尺度の中央の最初のスコアが測定の日井および床効果の低減によって変化に対する感度を改善させる。しかし、慢性疼痛に関する一部の試験では、疼痛強度の最低ラインを低く設定するか高く設定するかのいずれかであり、例えば、0～10のNRSで3[43]または5[25]と設定されている。個々の被験者の疼痛強度の最低ラインを4に設定すると、ほとんどの場合、神経障害性疼痛に関する試験では無作為割り付けされた患者のベースラインの群平均疼痛強度は5～7となるが[68]、OAに関する一部の試験では7～8の平均疼痛強度が観察されている[6]。

欧州医薬品庁[38]は、軽度の疼痛患者ではプラセボ反応率が高くなる可能性があるとの理論的根拠[参照 61,155]に基づき、神経障害性疼痛に関する臨床試験では中等度以上の疼痛を有する患者を組入れることを推奨している。大うつ病に関する臨床試験では、高い抑うつ重症度は治療効果が大きいことと関連していることを示すある程度のエビデンスがあるが[27,77,82]、ベースライン時の疼痛レベルが高ければ高いほど、治療効果が大きいかどうかについて検証した慢性疼痛試験は比較的少ない。有痛性DPNに関する3件の試験を対象とした最近の分析では、Zieglerら[157]は、ベースライン時の疼痛強度が0～10のNRSで6以上の被験者と無作為割り付けの最低要件であった4以上6未満の被験者の比較から、治療の絶対差が大きいことを指摘している。OAまたは関節リウマチでベースライン時の疼痛強度が高い患者は、ベースライン時の疼痛強度が低い患者よりも1クールの実薬治療による改善が大きいことも示されているが[11,12,15]、ベースライン時の疼痛強度が高い患者でプラセボまたは実薬対照との比較で高い有効性が認められるかどうかは不明である。

重要なことは、臨床試験の組み入れ基準が限られている場合(すなわち、要求される標的症状のレベルが高い)、患者が組み入れ基準を満たすようにベースライン時の症状重症度の評価の「インフレーション」が起こる可能性があることを示すある程度のエビデンスがある[79,82]。そのようなベースライン時の

インフレーションが研究実施施設の財政面などの検討事項や患者が新規治療を受ける機会を得たいという動機付けなどの因子を、どの程度反映しているかどうかは不明である。しかしながら、無作為割り付けに求められるベースライン時の疼痛強度を高く設定することによって慢性疼痛試験の分析感度を高めようとする全ての試みには、ベースライン時の疼痛強度の潜在的インフレーションを伴う可能性があり、このことによって分析感度が低下する可能性がある。慢性疼痛の臨床試験で、ベースライン時の疼痛の意図的または非意図的なインフレーションが起こることを想定して、以下の2つの方法のうち少なくとも1つにおいてその影響を軽減することが可能であろう：(1) ベースライン時の疼痛に関する特定の組み入れ基準に対して試験責任医師と患者の盲検化を維持する方法、(2) ベースライン時の疼痛とは無関係に全ての患者を無作為割り付けするが(これによってベースライン時のスコアを膨らませようとする金銭的および個人的なインセンティブが排除される)、主要有効性解析はベースライン時の疼痛が一定強度を上回る患者のみを検討することを予め定める方法。後者のアプローチを用いる場合は、ベースライン時の疼痛強度が事前に設定したカットオフ値を上回る被験者数を十分確保するために無作為割り付けで層別化の必要がある。しかし、これらのアプローチの追加的コストを分析感度に対する潜在的ベネフィットと比較検討する必要がある。

ベースライン時の疼痛強度が極めて高い患者(例 0～10のNRSで9以上)については、そのような極めて高いレベルの疼痛はおそらく心理社会的ストレスを反映したものであり、これらの患者が治療に反応する可能性は低いと考えられることから、除外を考慮することもある。さらに、例えば、ベースライン時の疼痛強度の評価が0～10のNRSで平均4以上であるが、疼痛日誌で3未満の評価が何日かあるようなベースライン時の疼痛強度の評価が極めて変化しやすい患者は除外できる。有痛性DPNに関する最近の3件の試験で、7日間のベースライン期間中の疼痛日誌の評価に25%を上回る変動がある患者を除外しているが[14,25,51]、これらのうち2件の試験では有意な治療効果が示されなかったため、このアプローチの潜在的ベネフィットは依然として不明である[14,51]。同様に、ベースライン時の疼痛強度の評価の変動が最小限または変動がなかった患者は疼痛変化を見分ける能力が限られていることもあり、これらの患者は臨床試験からの除外を考慮する必要がある。しかしながら、疼痛強度の極端な変動性(または安定性)は、根底にある病理生理学的機序の差を反映している可能性があり、そのような患者を除外すれば試験結果の一般化が限られるかもしれない。

ベースライン時の疼痛強度が鎮痛薬に関する臨床試験のアウトカムに及ぼす影響については、既存の臨床試験のデータベースをレトロスペクティブに分析することによって、例えば、ベースライン時の疼痛強度の変動の大きさが分析感度の改善と関連しているかどうかを判定することによってさらに理解が深まった。この疑問をプロスペクティブに検討するために、ベースライン時の異なる疼痛強度(または変動性)を有する被験者を対象に治療効果を比較する予め定めたサブグループ解析の実施が、今後の慢性疼痛の概念実証試験や確認臨床試験で可能と思われる。

慢性腰痛およびOAに関するRCTでは「フレアデザイン」がしばしば使用されており、このデザインには、全ての鎮痛薬を中止した後に患者の疼痛レベルおよび/またはそれに関連した障害がある程度増すに違いない患者、および事前に設定した基準を満たした患者のみが無作為割り付けされている[6,71]。この戦略によって、ベースライン時の疼痛強度が高く設定されるだけでなく、非ステロイド性抗炎症薬やオピオイド系鎮痛薬などの前治療効果のあった患者および同一の作用機序を有する別

の薬剤の効果がみられる可能性のある患者を特定できる。

3.5. 内科的および精神科的な併存疾患

慢性疼痛に関する臨床試験では多数の内科的および精神科的な併存疾患が除外基準とされている。これらは調査対象の治療および病状によって異なり、また被験者の安全性の確保、不均一性の軽減、研究結果の解釈に際してのあいまいさの回避を目的にプロトコールにしばしば含まれる。

慢性疼痛の患者は複数のタイプの疼痛を経験することが多く、慢性疼痛に関する RCT の診断基準を満たすことが他の疼痛症状の存在を排除することにはならない。研究対象症状以外の疼痛症状が存在すれば、特に標的疼痛症状に対する治療効果の評価に際して被験者の評価能力を妨げるかもしれない。例えば、痛性 DPN 患者は腰椎神経根障害も併発している可能性があり、FM 患者は慢性腰痛がある場合が多く、併存疾患に対して異なる治療効果があると標的疼痛の報告する際に混乱が起こる。したがって慢性疼痛試験では、治療対象の疼痛よりも重度の疼痛、同一部位の疼痛または治療対象疼痛と明確に区別できない疼痛を有する患者は除外されることが多い。さらに、標的疼痛を増強させる可能性があるその他の内科的症状、または評価を妨げる可能性がある症状、例えば PHN または DPN における感覚に影響を及ぼす皮膚症状[114]や OA 試験における通風の病歴[6]は除外できる。

慢性疼痛試験に組み入れられた被験者の安全性を確保し、発現する有害事象の解釈を促すために通常、追加的な臨床的除外基準が必要とされるが、これらは研究対象となる特定治療、ならびにそのリスク、薬理学、その他の治療との相互作用に関する知見によって異なる。そのような除外基準の例として、非ステロイド性抗炎症薬に関する試験での胃腸潰瘍の存在や、主に腎排泄により消失する薬剤に関する試験での低いクレアチニン・クリアランス値および腎機能不全の既往がある。

多数の慢性疼痛試験の除外基準として、精神障害、大うつ病、不安障害、自殺念慮または自殺企図、薬物乱用またはアルコール依存、認知機能障害などの神経精神医学的併存疾患がある。これらの併存疾患の既往歴がある、または現在これらの併存疾患がある人の参加は、少なくとも以下の3つの重大な懸念が考えられる：(1) 精神障害または認知機能障害のある人は患者の自己報告によるアウトカム指標の評価を正確に記入できない、(2) 自殺リスクのある人や意図的または偶発的に被験薬を過量摂取するリスクのある人、薬物乱用の既往歴のある人、および乱用の可能性や突然中止すると退薬症状の発現リスクのある治療薬を服用している人、(3) 抑うつ、不安、その他の心理社会的苦痛のある患者にプラセボ反応が増強され、また/または実薬治療反応が低下することによって臨床試験の分析感度が低下する可能性。単回投与試験の結果から、心理社会的苦痛が高い被験者はプラセボ反応が高く、オピオイド系鎮痛薬反応が低いため、苦痛が低い被験者よりも総合的な治療効果が小さくなることが示されている[149]。しかし、FM に関する最近の研究から、治療効果が抑うつ診断[120]や不安および抑うつ症状[4]によって異なることを示すエビデンスはない。しかし、抑うつや不安が分析感度に及ぼすマイナス影響があったとしても、FM 治療に用いられた薬剤がこれらの精神科的症状に及ぼす有益効果によって相殺された可能性がある[102]。

臨床試験の被験者の安全性を確保することは最優先事項であり、治療によって重篤な有害事象を引き起こすことが予想される併存疾患のある患者は除外すべきである。一般に、重症または重篤な有

害事象の発現率が高いと予想される患者を除外すると試験終了前の脱落率を低下させ、そうすることによって分析感度が高まる。しかしながら、臨床試験で除外基準数が多くなるほど、研究サンプルが抽出される地域社会の患者集団に対する得られた結果の一般化が低下し、よくみられる併存疾患があるような患者に対する治療リスク(およびベネフィット)についての学習経験が低下する。このように、安全性、分析感度、一般化の間には極めて重要なトレードオフ(相殺)が存在し、慢性疼痛に関する臨床試験へのこのような除外基準の使用を検討する際には極めて慎重に評価する必要がある[116]。

3.6. 併用薬およびレスキュー薬

被験者に対し安定用量の併用鎮痛薬の継続を認めるかどうか、これらの使用を除外するかどうか(特定の種類のレスキュー鎮痛薬を除く)についての問題は、慢性疼痛の確認臨床試験のデザインに重大な課題を提起している。併用鎮痛薬を被験者が継続できる臨床試験では、併用鎮痛薬が禁止されている試験に比べて、プラセボと比較した場合の追加的ベネフィットの証明の確率が低くなるとしばしば考えられている。一方、さまざまな慢性疼痛症状に適応される安全で有効な第一選択薬とみなされるような薬剤が増えつつある。これらの薬剤が広く利用できる地域では、いずれも服用していない患者やこれらの服用を簡単に中止すると思われる患者は一般に難治性であり、したがって新規の治療に効果を示す可能性が低くなると考えられる。このような患者を臨床試験に組入れると有効性が証明される可能性が低くなるかもしれない。さらに、慢性疼痛試験で他の鎮痛薬の併用を禁じれば、併用療法が一般的である日常の臨床診療の現場で得られた結果を一般化できる可能性が少なくなるかもしれない。

残念ながら、慢性疼痛試験の分析感度に及ぼす併用鎮痛薬の影響を評価したデータは限られている。4件の PHN 試験のメタアナリシスから、鎮痛薬を併用した被験者と併用しない被験者を対象としたプレガバリンの治療効果をプラセボと比較したところ同等であることが判明したが、これらは事後解析であった[28]。高濃度のカプサイシン・パッチに関する最近の RCT から、試験期間中に1種類以上の鎮痛薬の服用を継続した被験者と、そのような薬剤を服用しなかった被験者においてプラセボと比較して有意なベネフィットが認められた[8]。さまざまな併用鎮痛薬が安定用量で投与されている被験者とそうでない被験者を対象とした治療効果の比較を事前設定した解析により、併用薬が慢性疼痛試験に及ぼす影響への理解が深まった[8,132]。

治療効果の検出が改善すると考えられる理由やその他の理由(例えば、忍容性および安全性の評価の解釈が容易となる)により併用鎮痛薬を除外すれば、プラセボ群の倫理的利用は通常、レスキュー鎮痛薬の使用を必要とし、これによって許容できない疼痛を経験する被験者にも試験継続が可能になる。例えば、神経障害性疼痛に関する多数の試験では、アセトアミノフェン(パラセタモール)がレスキュー鎮痛薬(最高用量は2~4 g/日が多い)として認められている。このようなレスキュー鎮痛薬が許容される場合、被験者が使用した量を記録および報告することが重要である。有効治療に関する試験のプラセボ群ではレスキュー鎮痛薬の量は多くなる可能性があり、データの分析および解釈ではこの点を考慮すべきである。

アセトアミノフェンなどのレスキュー薬は低用量であっても、レスキュー薬の薬理効果およびプラセボ効果は実薬治療群とプラセボ群の間での被験者反応の差を弱める可能性がある。被験者がレスキ

ユー薬を必要とするような疼痛が発現した場合、すみやかに治療を中止することを奨励しているならば、倫理的な面からレスキュー薬を認めないということもあり得るかもしれない。そのような試験は、耐え難い疼痛による中止までの期間に実薬群とプラセボ群の間で差を検討するように設計できる。レスキュー薬の禁止に対してそのようなアプローチを用いた慢性疼痛試験は少ないことから、確認試験でそのような試験デザインを考慮する前に追加的研究が必要となるであろう。

併用鎮痛薬(レスキュー薬以外のもの)を禁じる場合、ベースライン期間の開始前に除外対象の薬剤を安全に中止すると共に、それらの薬剤の影響(すべての退薬症状または持続中の治療効果)の停止が確保できるよう十分な休薬期間がなければならない。これによってベースラインのデータが薬剤中止の一過性の影響を受けないようにすることができる。このような休薬のために必要とされる期間は、除外対象の薬剤の薬物動態学および薬力学的な特徴によって大きく異なる。

併用鎮痛薬を認める場合、被験薬および市販薬を含む許可された薬剤すべての薬物動態および薬力学を徹底的に評価することにより潜在的な相互作用を明らかにしなければならない。被験薬に対する禁忌薬は全て除外すべきであり、全ての併用鎮痛薬は組入れ前の一定期間および試験期間中に安定用量を維持する必要があり、医学的に必要となる場合を除いて試験期間中に中止してはならない。さらに、試験開始後は鎮痛薬による治療を開始してはならない。これは慢性疼痛患者が用いることが多いその他の疼痛療法(例えば、理学療法、心理社会的介入、侵襲性介入)にも適用される。特定状況の下では、試験期間中における安定用量の併用鎮痛薬使用が治療反応に及ぼす影響をベースライン共変量として、または併用鎮痛薬を投与された被験者と投与されていない被験者を比較するサブグループ解析による有効性アウトカムの統計解析によって検証できる[8,58]。

慢性疼痛試験において併用鎮痛薬を認めるための異なるアプローチがある。1 つのアプローチは、全ての鎮痛薬について試験期間を通して安定用量の継続を認めるというものである。2 つ目は、既存治療薬の安定用量に被験薬を「上乘せ」して検討するものである。規制当局への申請に関連して実施される試験について、最もわかりやすいアプローチは、研究対象の特定の慢性疼痛症状に対して国または地域当局から承認されている既存の治療を用いることである。ただし、特定の慢性疼痛症状に対して規制当局から承認されていないが、一般に有効と考えられている薬剤(例えば、FM または神経障害性疼痛に対する三環系抗うつ薬)があるが、承認されている適応は当該規制当局によって異なる場合が多い。したがって、3 つ目のアプローチは、限られた数の特定の鎮痛薬(例えば、規制当局が承認したもの、または有効性のエビデンスが一般に受け入れられているもの)のみを、試験期間を通して安定用量で使用することを認め、それぞれについて最高用量を明記するというものである。例えば、有痛性 DPN 患者を対象とした試験では、以下の薬剤のうち 2 つに限って安定用量で使用を認めることができるかもしれない:(1) ガバペンチンまたはプレガバリン、(2) オピオイド系鎮痛薬またはトラマドール、または(3) 三環系抗うつ薬または選択的セロトニン・ノルエピネフリン再取り込み阻害薬。もちろん、使用はこれらの 3 種類の薬剤のうち 1 種類のみに限定するか 3 種類すべてとし、有効用量を最小限、適度または最大限に反映できる最高用量を事前に設定することができる[11]。しかし、併用鎮痛薬の許可に関してこのような構造化されたアプローチを用いている臨床試験はわずかであることから、このような意思決定の基となっているデータは少ない[132]。

慢性疼痛に関する確認試験の設計では、レスキュー薬の持続期間が限られており、理想的には主

要有効性アウトカムと一致しないことを条件として、被験者に無関係な急性疼痛症状(例えば、歯科治療や外傷)が発現した場合にレスキュー薬の許可を考慮することができる。

3.7. 前治療に対する反応

現在、作用機序が異なる多種多様な薬剤がさまざまな慢性疼痛症状に適応されており、有効性に関するエビデンスが広く受け入れられているが、十分でない鎮痛効果、忍容性不良または不十分な用量漸増によって地域社会で相当数の患者で十分な治療効果がみられていない[53]。このような患者が難治性の疼痛を有している場合や一般に薬剤の副作用に敏感である場合、これらの患者を臨床試験に組入れると分析感度が低下し、有害事象例や中止例の数が増す可能性がある。

これらの理由により、新規の作用機序を有する薬剤がこのような患者に有効となる可能性を知ることが重要であるが、有効性欠如または忍容性不良により複数の前治療に効果がなかった患者の除外を考慮できる。レトロスペクティブな患者報告に基づいて有効性欠如および忍容性不良を明確にすることは可能であるが[136]、このような判定を行うことには問題が多く、薬剤用量や治療期間が十分に思い出せないことから限られる。いくつかの臨床試験の解析は、研究開始前の種々の鎮痛薬使用歴があるサブグループで治療効果を比較することにより前治療の役割を検討している[例 71]。慢性疼痛に関する臨床試験で前治療における治療反応が分析感度に及ぼす影響への知見をさらに認識するため、被験者の治療歴に関して十分な情報が収集されている終了している試験を対象としたレトロスペクティブな評価を実施すべきである。

4. 慢性疼痛臨床試験の相

慢性疼痛の臨床試験のために患者をスクリーニングし、十分な説明に基づき患者からの同意を得た後に疼痛などの変数が評価されるのが一般的であり、続いて被験者を選択し、種々の治療群に無作為割り付けする。RCT では無作為割り付け前の期間が用いられることが多く、これらの期間は被験者への治療(無治療、プラセボ、被験薬)、盲検化の程度(非盲検、被験者を対象とした単盲検、被験者と試験担当者を対象とした二重盲検)および導入期間の目的によって異なる[101]。上述したように、このような期間は被験者が無治療または併用鎮痛薬の安定用量を投与されている期間の疼痛の重症度およびばらつきを確認するために用いられる。無作為割り付け前の期間は、被験者がプラセボで鎮痛を経験するかどうか、実薬で鎮痛や耐え難い副作用を経験するかどうかを評価するためにも用いられる。さらに、無作為割り付け前の期間は研究手順に従わない被験者を特定した後に除外するためにも利用できる[83]。例えば、最低日数分(例 7 日中 4 日または 5 日)の疼痛日誌の記入を完了しない被験者は、このような早期段階のノンアドヒアランスは無作為割り付け後のその後の研究手順に対するノンアドヒアランスを予測するという想定に基づいて、無作為割り付けから除外できる[7,118]。

4.1. ベースライン期間

最も簡単なタイプのベースライン期間において、被験者にはプラセボも実薬も投与されず、無作為割り付けへの適格性を確認し、その後の治療効果を評価するためのベースラインとするためにデータ