

分子薬理学研究における
臨床副作用自発報告ビッグデータの活用

2019

長島 卓也

謹呈

様

拙書ではございますが、御一読頂ければ幸いに存じます。
今後ともご指導・ご鞭撻を賜りますようお願い申し上げます。

令和2年3月23日

長島 卓也

分子薬理学研究における
臨床副作用自発報告ビッグデータの活用

2019

長島 卓也

目次

緒言.....	1
第一章 非定型抗精神病薬による高血糖に対する解決策の提示 及び分子機序の探索.....	3
序論.....	3
方法.....	4
結果.....	9
考察.....	23
第二章 分子薬理学研究ツールとしての有害事象ビッグデータ の解析手法開発.....	26
序論.....	26
方法.....	27
結果.....	28
考察.....	44
総括及び結論.....	46
謝辞.....	47
発表論文目録.....	48
参考文献.....	49

緒言

近年、医療・創薬における現状の様々な問題を根底的に革新する先進的な戦略として「ビッグデータ」の利活用に対する期待が高まっている。例えば、医薬品による副作用・有害事象は医療従事者・企業を介して FDA や PMDA といった規制当局に報告されるが、現在ではこのような実臨床データが匿名化された形で Web 上に公開されており、1000 万件以上の患者データを二次的に解析することが可能である。本研究では、このような有害事象ビッグデータを分子薬理学研究分野においてどのように活用できるかについて検討を行い、以下の新知見を得た。

第一章では、有害事象ビッグデータ解析の活用例として非定型抗精神病薬(クエチアピン)使用に伴う高血糖・糖尿病に注目した。有害事象データベースの薬剤疫学的解析により、ビタミン D 製剤の併用という新たな解決策を見出した。また、この仮説を動物・細胞を用いた分子薬理学的実験により実証した。更に、クエチアピン・糖尿病・ビタミン D という 3 項の共通点から糖尿病症状における *Pik3r1* 遺伝子を介した新たな分子メカニズムの関与を明らかにした。

第二章では、分子薬理学研究ツールとしての有害事象ビッグデータの解析手法開発を試みた。データの表現方法として、分割表データを分子薬理学実験と同様に平均値、標準誤差及び N に要約し可視化できることを示した。また背景の調整方法として、傾向スコアマッチング法により分子薬理学実験と同じく N と背景因子を統一できることを示した。更に群間の比較方法として、差を指標とする Score 検定法が分子薬理学研究に応用する上で有用であることを示した。

以上、本研究は、臨床副作用自発報告ビッグデータが、薬物有害事象の新たな解決策の提示、既存薬の新たな薬効の探索、病態の新たな分子メカニズムの解明に有用であることを明らかにしたものであり、本手法を分子薬理学研究に応用するために必要な技術に関する知見を提供したものである。

以下にこれらの研究成果について論述する。

なお、本文及び図中で使用した略語は以下の通りである。

2-NBDG	2-deoxy-2-[(7-nitro-2,1,3-benzoxadiazol-4-yl)amino]-D-glucose
AMPK	adenosine monophosphate-activated protein kinase
ANOVA	analysis of variance
cDNA	complementary deoxyribonucleic acid
CI	confidence interval
DM	diabetes mellitus
DMEM	Dulbecco's modified Eagle medium
FAERS	FDA Adverse Event Reporting System
FBS	fetal bovine serum
FDA	Food and Drug Administration
GEO	Gene Expression Omnibus
GLM	generalized linear model
GLUT4	glucose transporter type 4
HS	horse serum
Insr	insulin receptor
Irs1	insulin receptor substrate 1
JADER	Japanese Adverse Drug Event Report database
KEGG	Kyoto Encyclopedia of Genes and Genomes
KRPB	Krebs Ringer phosphate HEPES
LRT	likelihood ratio test
mRNA	messenger ribonucleic acid
OR	odds ratio
PI3K	phosphatidylinositol 3-kinase
Pik3r1	phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 1
Pik3r2	phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 2
PMDA	Pharmaceuticals and Medical Devices Agency
qRT-PCR	quantitative reverse transcription-polymerase chain reaction
rRNA	ribosomal ribonucleic acid
Slc2a4	solute carrier family 2 (facilitated glucose transporter), member 4

第一章

非定型抗精神病薬による高血糖に対する 解決策の提示及び分子機序の探索

序論

クエチアピンやオランザピンに代表される非定型抗精神病薬は統合失調症、双極性障害、うつ病といった様々な精神疾患に対し広く用いられているが、有害事象として血糖上昇との関与が知られており、臨床的に大きな問題となっている[1,2]. 非定型抗精神病薬による高血糖は糖尿病の発症にとどまらず、糖尿病性ケトアシドーシスや糖尿病性昏睡といった重篤な症状を引き起こし、中には死亡例も報告されている[3].

非定型抗精神病薬と高血糖・糖尿病との関与は広く知られているにもかかわらず、その機序についてはほとんど解明されていない。これまで報告されている機序としては、体重増加、膵臓 β 細胞からのインスリン分泌の減少、インスリン抵抗性、レプチン作用障害等が挙げられている[4]. メトホルミンのような抗糖尿病薬が有効であるとするという臨床研究も存在するが、抗精神病薬による高血糖には複数の要因が関与してくるため単剤療法は困難であり[5], 更なる治療・予防法が望まれている。

このような状況を打開するために、著者は有害事象ビッグデータの活用可能性に着目した。米国のFAERSデータベースには世界中の有害事象報告が蓄積されており、有害事象発生時の患者の年齢、性別、原疾患、併用薬といった情報を活用することができ、FAERSの解析を通じて未知の薬物間相互作用の同定に成功した報告も複数存在する[6-8]. そこで著者は、FAERSの解析を通じて非定型抗精神病薬による高血糖を抑制しうる併用薬を探索し、そこで得られた仮説を動物・細胞を用いた実験検証と組み合わせることで、分子薬理学分野における有害事象ビッグデータ解析の有用性について検討を行った。

方法

FAERS データベースの解析

FAERS のデータは FDA のウェブサイト (<http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/>) から入手し、2004 年第 1 四半期(04Q1)から 2014 年第 2 四半期(14Q2)までの 5,821,345 件の報告から、FDA の推奨に従い最新の症例番号を有する報告のみに絞り込み重複を除いた 4,547,841 件の症例について解析した。また、FAERS に登録されている薬物名は任意記載形式であるため、ライフサイエンス辞書シソーラスを駆使したテキストマイニングにより薬物名を一般名に統一した。

各薬物と糖尿病報告との相関はオッズ比(OR)及び 95%信頼区間(CI)により評価した[34]。FAERS の各症例は以下のように 4 つのグループに分割した: (a) 興味のある薬物(例:クエチアピン, ビタミン D)を使用し、かつ有害事象として糖尿病が報告されている症例, (b) 興味のある薬物を使用し有害事象として糖尿病が報告されていない症例, (c) 興味のある薬物を使用しておらず有害事象として糖尿病が報告されている症例, (d) 興味のある薬物を使用しておらず有害事象として糖尿病も報告されていない症例。これらの値について、オッズ比及びその 95% 信頼区間は以下のように計算した:

$$OR = \frac{a/b}{c/d}, 95\% CI = \exp\left(\log(OR) \pm 1.96 \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}\right)$$

ここで、 a, b, c, d は各グループの症例数であり、 \log は自然対数である。オッズ比の 95% 信頼区間下限が 1 を上回る場合、興味のある薬物(例:クエチアピン)の使用と糖尿病報告割合の増加との間に統計学的に有意な相関があると解釈した。反対に、オッズ比の 95% 信頼区間上限が 1 を下回る場合、興味のある薬物(例:ビタミン D)の使用と糖尿病報告割合減少との間に統計学的に有意な相関があると解釈した。

「糖尿病」は表 I-S1 の 17 種類のいずれかの有害事象名の記載により定義した。また、「ビタミン D 製剤」は表 I-S2 の 15 種類の薬物名により定義した。原疾患情報としては薬物の適応症情報を用いた。オッズ比はサンプルサイズが小さい場合に変動しやすいため、クエチアピン使用患者における糖尿病を低減しうる併用薬を探索する際には併用症例数が 1000 人以上存在する併用薬のみに解析対象を絞り込んだ。

Adverse event name	
Diabetes mellitus	Insulin-requiring type 2 diabetes mellitus
Type 2 diabetes mellitus	Diabetic hyperglycaemic coma
Diabetes mellitus inadequate control	Diabetic ketoacidotic hyperglycaemic coma
Diabetic ketoacidosis	Insulin-requiring type II diabetes mellitus
Diabetic coma	Diabetes with hyperosmolarity
Type 1 diabetes mellitus	Insulin resistant diabetes
Diabetes mellitus non-insulin-dependent	Latent autoimmune diabetes in adults
Diabetes mellitus insulin-dependent	Fulminant type 1 diabetes mellitus
Diabetic hyperosmolar coma	

表 I-S1. FAERS 解析における「糖尿病」の検索名

Resolved drug name	
Vitamin D	Calcifediol
Vitamin D-calcium combination	Maxacalcitol
Cholecalciferol	Eldecalcitol
Paricalcitol	Tacalcitol
1-Hydroxycholecalciferol	Falecalcitriol
Calcitriol	Cholecalciferol-calcium combination
Ergocalciferol	Dihydroxycholecalciferol
Doxercalciferol	

表 I-S2. FAERS 解析における「ビタミン D 製剤」の検索名

GEO 及び KEGG PATHWAY データベースを用いた分子メカニズムの解析

非定型抗精神病薬による高血糖の分子メカニズム解明には KEGG PATHWAY データベース[9,10]と DrugMatrix[11]に含まれるマイクロアレイデータとを組み合わせることにより予測した。DrugMatrix のデータは GEO データベース[12,13]に登録されており、クエチアピン(500 mg/kg)または対照溶液(カルボキシメチルセルロース)を5日間経口投与したラット肝臓の遺伝子発現マイクロアレイデータを入手した(GEO Accession No. GSE 59923; <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE59923>)。得られた遺伝子発現データのうち、KEGG PATHWAY データベースの“insulin resistance”経路(Entry No. map04931; http://www.kegg.jp/dbget-bin/www_bget?pathway+map04931)に含まれる遺伝子の発現変動を網羅的に解析し、骨格筋におけるインスリン抵抗性経路へマッピングした。

実験動物

実験には5-6週齢の雄性 ICR マウス(日本 SLC)を使用し、室温 24±1°C、湿度 55±10%、12時間/12時間の明暗周期条件下で飼育した。餌及び水はグルコース負荷試験における絶食期間を除き自由摂取のもとで飼育した。実験は全て京都大学動物実験委員会による審査・承認を受け、「動物実験に関する日本薬理学会指針」及び「京都大学における動物実験の実施に関する規定」を遵守して実施した。

薬物・試薬

In vivo 実験において、クエチアピン(和光純薬工業)及びメタンフェタミン(大日本住友製薬)は1% Tween 80 (ナカライテスク)溶液に、D-グルコース(ナカライテスク)は純水に溶解させてから投与した。クエチアピン、メタンフェタミン、D-グルコースの各薬物は10 ml/kg body weight となるよう溶液を腹腔内に投与した。

ビタミン D₃ (コレカルシフェロール)は餌投与とした。普通餌には1 kg 当たり 1370 IU ビタミン D₃ を含む MF 飼料(オリエンタル酵母)を用いた(マウスが1日に6 g の飼料を摂取すると仮定して8 IU ビタミン D₃/day)。コレカルシフェロール添加餌は MF 飼料に 200,000 IU ビタミン D₃/kg を混合して作製した(1200 IU ビタミン D₃/day)。

In vitro 実験では上記以外の試薬としてインスリン(Biological Industries), カルシトリオール(Toronto Research Chemicals), 2-NBDG, LY294002 及び compound C

(Cayman Chemical), DMEM 及び FBS (Sigma-Aldrich), そして HS (Invitrogen)を用いた.

グルコース負荷試験

マウスにビタミン D (コレカルシフェロール)添加餌又は普通餌を 1 週間投与した後, 16 時間絶食させ, 150 分間のグルコース負荷試験を実施した. 0 分時点でクエチアピン(10 mg/kg)または対照溶液(1% Tween 80)を腹腔内投与し, その 30 分後に D-グルコース(3 g/kg)を腹腔内投与した. 各タイムポイントにおいて血液サンプル(5 µl)を尾静脈から回収し, 血糖値は Accu-Chek Blood Glucose Meter (Roche Diagnostics)を, 血中インスリン濃度は Mouse Insulin Enzyme-Linked Immunosorbent Assay Kit (森永生化学研究所)を用いて測定した.

走行距離(ロコモーター)評価試験

雄性 ICR マウスに普通餌又はビタミン D 添加餌を 1 週間投与した後, 対照溶液(1% Tween 80), メタンフェタミン(2 mg/kg), メタンフェタミン(2 mg/kg) + クエチアピン(10 mg/kg)混合溶液のいずれかを腹腔内投与し, その 30 分後にオープンフィールド(50 cm × 50 cm × 50 cm)内へ移動させ, 走行距離を 30 分間測定した.

定量的逆転写ポリメラーゼ連鎖反応

グルコース負荷試験の終了直後にマウスを頸椎脱臼し, 右大腿部の筋組織サンプルを回収した. mRNA は ISOGEN Reagent (ニッポンジーン)を用いて抽出し, 抽出した mRNA (0.5 µg)を ReverTra Ace qPCR RT Kit (東洋紡)により cDNA へ逆転写した. 定量的逆転写ポリメラーゼ連鎖反応(qRT-PCR)は StepOne Real-Time PCR System (Life Technologies)及び THUNDERBIRD SYBR qPCR Mix (東洋紡)を用いて実施した. 内部標準遺伝子には *18s rRNA* を用い, mRNA 発現量データは対照群に対する相対的発現量で表した. 実験に用いたプライマーの配列は以下の通りである: *18s rRNA*: 5'-GTA ACC CGT TGA ACC CCA TT-3' & 5'-CCA TCC AAT CGG TAG TAG CG-3', *Insr*: 5'-CCC AGG CCA TCC CGA AAG-3' & 5'-TCT CAA ATG GCC TGT GCT CC-3', *Irs1*: 5'-TTA GGC AGC AAT GAG GGC AA-3' & 5'-TCT TCA TTC TGC TGT GAT GTC CA-3', *Pik3r1*: 5'-GAC AGC GAA GCG ACG GC-3' & 5'-GTC TGA TTT TAC TGC CAC GCT C-3', *Akt2*: 5'-CGC TTG CGG TCT GAT GTT TT-3' & 5'-AAT ACC GCC TTT TCC AGC CA-3', *Slc2a4*:

5'-TTA TTG CAG CGC CTG AGT CT-3' & 5'-GGG TTC CCC ATC GTC AGA G-3'.

細胞培養及び分化

マウス骨格筋由来 C2C12 細胞株は 100 mm ディッシュ上で 10% FBS 含有 DMEM 培地を用い 37°C, 5% CO₂ 条件下で培養した. 96 ウェル黒プレートに播種した翌日に, 培地を 2% HS 含有 DMEM に変更し筋管細胞へと分化させた. 実験には分化から 3-5 日後の細胞を用いた.

グルコース取込試験

グルコース取込試験はグルコースの蛍光誘導体である 2-NBDG を用いて実施した. 96 ウェル黒プレート上で培養された C2C12 筋管細胞に対し, 各薬液を処置後, インスリン(1 μ M)含有/非含有 KRPH 緩衝液(136 mM NaCl, 4.7 mM KCl, 1 mM MgSO₄, 1 mM CaCl₂, 5 mM KH₂PO₄, 20 mM HEPES)で 15 分間処置し, 続いて 2-NBDG (50 μ M)溶液を 20 分間処置した. 遊離 2-NBDG は KRPH バッファーで 3 回洗浄して除去し, 細胞に取り込まれた 2-NBDG 蛍光は励起波長 480 nm, 放出波長 540 nm として蛍光マイクロプレートリーダー(FDSS/ μ Cell; 浜松ホトニクス)により測定した. グルコース取込量はインスリン非処置群に対する蛍光強度の比で算出した.

統計解析

FAERS 及び GEO データベースの統計解析は R version 3.1.2 Software (R foundation for Statistical Computing)を用いた. 動物・細胞実験で得られたデータは Prism 5 Software (GraphPad)により解析した. 2 群間の平均値の差の検定には unpaired two-tailed *t* test with Welch's correction を用い, 3 群以上の比較には Kruskal-Wallis test 及び Dunn's *post hoc* multiple comparison test 又は one-way ANOVA 及び Tukey's *post hoc* multiple comparison test を用いた. 時系列データについては two-way ANOVA for repeated measures 及び Bonferroni's *post hoc* multiple comparison test を用いた. $P < 0.05$ の場合に統計学的に有意な差があると解釈した.

結果

非定型抗精神病薬の使用と糖尿病報告との相関解析

著者はまず、非定型抗精神病薬の使用と糖尿病報告との相関について FAERS データベースの解析により検討した(表 I-1)。FAERS データベースの不均衡解析の結果、非定型抗精神病薬の使用と糖尿病報告割合の増加との間に有意な相関が認められた。非定型抗精神病薬同士で比較すると、クエチアピンとオランザピンについては糖尿病報告割合増加との間に強い相関(OR > 20)が、リスペリドン、アリピプラゾール及びジプラシドンについては中程度の相関(5 < OR < 20)が、クロザピンについては弱い相関(OR < 5)が認められた。各薬物により程度の差はあるものの、非定型抗精神病薬により糖尿病リスクが増加するという臨床上の知見に一致する結果が本解析により得られた。著者はこれらの薬物のうち、糖尿病のオッズ比が高く、かつ報告数の多いクエチアピンに着目して後の検討を進めることにした。

Drug A	DM without drug A (%)	DM with drug A (%)	Odds ratio (95% CI)
Quetiapine	37373/4482388 (0.83)	10809/65453 (16.51)	23.53 (22.99–24.07)
Olanzapine	41668/4515358 (0.92)	6514/32483 (20.05)	26.93 (26.17–27.72)
Risperidone	44382/4513579 (0.98)	3800/34262 (11.09)	12.56 (12.13–13.01)
Aripiprazole	46007/4518947 (1.02)	2175/28894 (7.53)	7.91 (7.57–8.28)
Ziprasidone	46569/4534867 (1.03)	1613/12974 (12.43)	13.68 (12.98–14.43)
Clozapine	47124/4517820 (1.04)	1058/30021 (3.52)	3.47 (3.26–3.69)

表 I-1. FAERS データベースにおける非定型抗精神病薬(drug A)と糖尿病報告との相関解析

代表的な非定型抗精神病薬同士での比較。糖尿病の検索用語は表 I-S1 に示す。

クエチアピン使用患者の糖尿病報告に対する 有害事象名の定義、年齢、性別、原疾患、併用薬の影響

続いて著者は、クエチアピン使用に伴う糖尿病報告に対する有害事象名の定義、年齢、性別、原疾患、併用薬の影響について検討した(表 I-2)。クエチアピンの使用と糖尿病報告割合増加との相関は代表的な有害事象名(例: “diabetes mellitus”, “diabetic ketoacidosis”, “diabetic coma”)のみで糖尿病を定義しても依然として認められた。各年齢層に分割して解析したところ、オッズ比は40–49歳で最も高く、40歳未満、50–59歳、60–69歳、70歳以上の順に低下していくという結果が得られ、これは先行研究における知見[2]と一致する結果であった。性別についても同様に解析したところ、男女ともに同程度のオッズ比が認められた。また、統合失調症や双極性障害といった代表的な原疾患を有する患者層や、ハロペリドールやリチウムといった標準的治療薬を用いている患者層に絞って解析しても、依然として高いオッズ比が認められた。以上の結果から、クエチアピン使用に伴う糖尿病報告の増加は患者の年齢、性別、原疾患、併用薬にかかわらず認められることが示唆された。

Group	DM without quetiapine (%)	DM with quetiapine (%)	Odds ratio (95% CI)
Limited by an adverse event name			
Diabetes mellitus	20092/4482388 (0.45)	7668/65453 (11.72)	29.47 (28.67–30.30)
Diabetic ketoacidosis	2897/4482388 (0.06)	1258/65453 (1.92)	30.30 (28.35–32.39)
Diabetic coma	885/4482388 (0.02)	876/65453 (1.34)	68.69 (62.55–75.44)
Stratified by age			
<40 years	3622/657664 (0.55)	2229/16317 (13.66)	28.57 (27.03–30.20)
40–49 years	3982/391629 (1.02)	2813/11828 (23.78)	30.38 (28.82–32.02)
50–59 years	5929/547124 (1.08)	1690/10574 (15.98)	17.36 (16.39–18.40)
60–69 years	5170/536481 (0.96)	409/4932 (8.29)	9.29 (8.37–10.32)
≥70 years	4100/616472 (0.67)	175/5067 (3.45)	5.34 (4.58–6.23)
Stratified by gender			
Male	14129/1572620 (0.90)	4512/24882 (18.13)	24.43 (23.56–25.33)
Female	18996/2537316 (0.75)	6010/37798 (15.90)	25.06 (24.30–25.85)
Stratified by primary disease			
Schizophrenia	1060/27277 (3.89)	1398/4160 (33.61)	12.52 (11.45–13.68)
Bipolar disorder	649/23332 (2.78)	2459/10676 (23.03)	10.46 (9.56–11.45)
Stratified by drug combination			
Haloperidol	308/12393 (2.49)	1219/2712 (44.95)	32.04 (27.96–36.71)
Lithium	351/13116 (2.68)	785/3771 (20.82)	9.56 (8.38–10.91)

表 I-2. クエチアピン使用に伴う糖尿病報告に影響を与える各因子の詳細検討

クエチアピン使用患者における糖尿病を抑制しうる併用薬の網羅的探索

続いて著者は、クエチアピン使用による糖尿病を抑制しうる併用薬を FAERS データベースの解析により網羅的に探索した(表 I-3)。クエチアピン使用症例に解析対象を絞り込んだ上で不均衡解析を実施したところ、プレガバリン、ラモトリギン、ビタミン D 製剤といった薬物が糖尿病報告割合の低下と有意に相関していることが見出された。このうち、最もオッズ比の小さい薬物はビタミン D 製剤であった。オランザピン使用症例に解析対象を絞り込んだ上で不均衡解析を実施しても同様の結果が得られた(OR = 0.21, 95% CI: 0.14–0.32)。一方で、ビタミン D 製剤自体には糖尿病報告割合の変化との間に統計学的に有意な相関は認められなかった(OR = 1.01, 95% CI: 0.94–1.07, $P = 0.86$)。これらの結果から、ビタミン D 製剤はクエチアピンやオランザピンといった非定型抗精神病薬と相互作用し高血糖リスクを抑制しうるという仮説が得られた。プレガバリン(OR = 1.33, 95% CI: 1.25–1.41)やラモトリギン(OR = 1.41, 95% CI: 1.30–1.52)といった他の候補薬物は、それ自体が僅かながらも糖尿病報告割合の増加と有意な相関が認められたため、以降の検討では候補併用薬としてビタミン D 製剤に着目した。

Drug B	DM without drug B (%)	DM with drug B (%)	Odds ratio (95% CI)
Pregabalin	10625/64076 (16.58)	184/1377 (13.36)	0.78 (0.66–0.91)
Lamotrigine	10387/61398 (16.92)	422/4055 (10.41)	0.57 (0.51–0.63)
Vitamin D	10743/64282 (16.71)	66/1171 (5.64)	0.30 (0.23–0.38)

表 I-3. FAERS データベースにおいてクエチアピン(drug A)使用症例の糖尿病報告割合の低下と相関の認められる併用薬(drug B)の網羅的探索

解析対象をクエチアピン使用症例のみに絞り込んだ上で不均衡解析を実施し、オッズ比の小さい併用薬上位 3 件を抽出した。ビタミン D 製剤の検索用語は表 I-S2 に示す。

グルコース負荷試験における血糖値・血中インスリン濃度に対する クエチアピン・ビタミン D 投与の影響

続いて著者は、クエチアピンによる血糖上昇がビタミン D の併用により抑制されるという仮説を薬理学的実験により検証した。雄性 ICR マウスに普通餌又はコレカルシフェロール添加餌(1200 IU vitamin D₃/day)を 1 週間投与しても、体重に有意な変化は認められなかった(図 I-S1)。16 時間の絶食後、グルコース負荷試験を実施したところ、普通餌飼育クエチアピン(10 mg/kg)投与群では普通餌飼育対照溶媒投与群に比べ有意な血糖上昇及び血中インスリン濃度の上昇が認められた(図 I-1)。一方で、コレカルシフェロール添加餌飼育クエチアピン投与群ではこれらの上昇に対する有意な抑制が認められ、FAERS の解析で得られた仮説に一致する結果が動物実験レベルで確認された(図 I-1)。また、クエチアピンの統合失調症治療作用について走行距離(ロコモーター)評価試験により検討したところ、コレカルシフェロール添加餌投与による影響は認められなかったことから、ビタミン D を併用してもクエチアピンの主作用は減弱されないと考えられた(図 I-S2)。高インスリン血症はインスリン抵抗性状態の特徴の 1 つであることから[14]、クエチアピンはインスリン抵抗性を引き起こしており、このインスリン抵抗性がビタミン D の併用により改善されるという可能性が示唆された。

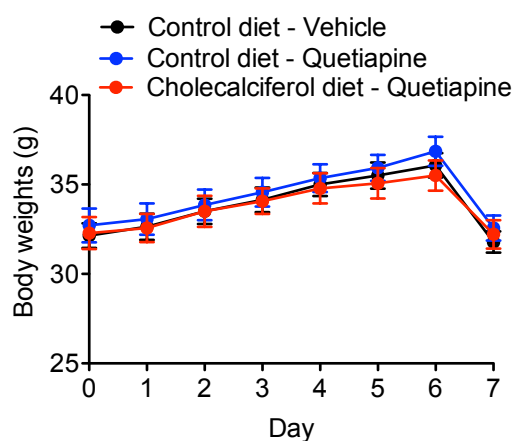


図 I-S1. 体重に対する普通餌及びコレカルシフェロール添加餌投与の影響

ICR マウスを普通餌又はコレカルシフェロール添加餌(1200 IU vitamin D₃/day)で 1 週間飼育し、16 時間絶食させた時の体重変動を測定した。データは平均値±標準誤差で示している。N=14。

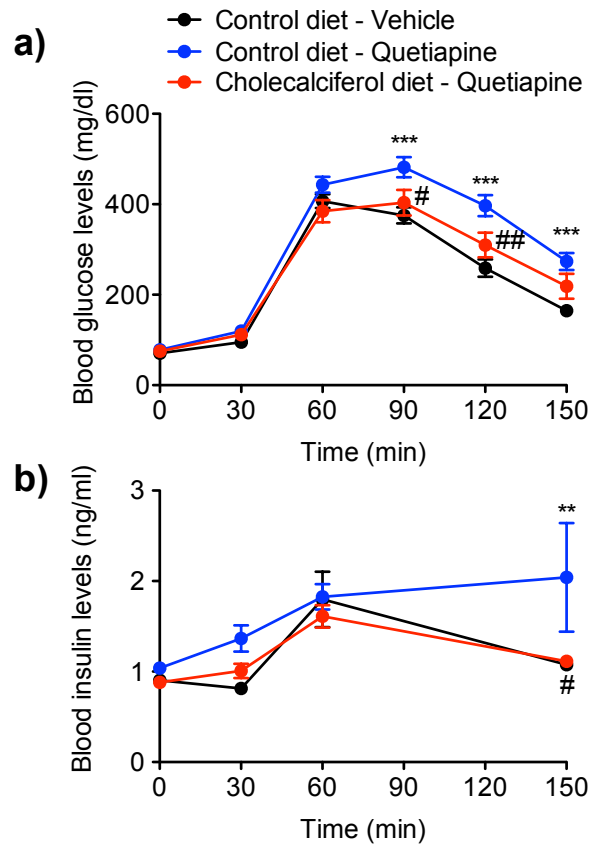


図 I-1. クエチアピンによる血中グルコース及びインスリン濃度の上昇に対するコレカルシフェロール添加餌投与の影響
 0分時点でクエチアピン(10 mg/kg)又は対照溶媒を, 30分時点でグルコース(3 g/kg)を腹腔内に投与した. (a)血糖値($N = 14$)及び(b)血中インスリン濃度($N = 8$)を図に示すタイムポイントで測定した. データは平均値 \pm 標準誤差で示している. ** $P < 0.01$, *** $P < 0.001$ vs. 対照群; # $P < 0.05$, ## $P < 0.01$ vs. クエチアピン投与群.

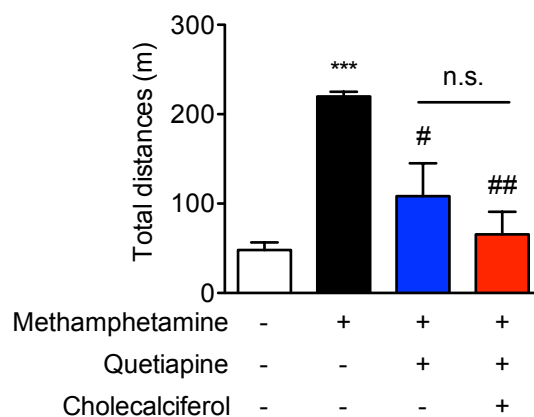


図 I-S2. クエチアピンの統合失調症治療作用に対するコレカルシフェロール添加餌投与の影響

メタンフェタミンを投与されたマウスは走り回るようになり、これはドパミン D₂ 受容体の刺激を介した統合失調症の陽性症状モデルと考えられている[35]. 著者はこのモデルを用いてクエチアピンの主作用に対するコレカルシフェロール添加餌投与の影響を検討した. メタンフェタミン(2 mg/kg)投与群における総走行距離の増加はクエチアピン(10 mg/kg)投与群において有意に抑制された. このクエチアピン投与による統合失調症モデルの改善作用はコレカルシフェロール添加餌併用群においても認められ、普通餌飼育クエチアピン投与群との間に有意な差は認められなかった. バーは平均値及び標準誤差を示している. $N=4$, *** $P < 0.001$ vs. 対照群; # $P < 0.05$, ## $P < 0.01$ vs. メタンフェタミン投与群; n.s. not significant.

GEO 及び KEGG PATHWAY データベースの統合解析による クエチアピン誘発インスリン抵抗性の分子メカニズム探索

続いて著者は、クエチアピンによるインスリン抵抗性の分子メカニズムを探索するため、GEO データベースに含まれるマイクロアレイデータ (Accession No. GSE59923) を利用し、KEGG PATHWAY データベースのインスリン抵抗性経路 (Entry No. map04931) に含まれる遺伝子の発現について網羅的に解析・検討した (図 I-2)。その結果、クエチアピンの投与により発現変動するインスリン抵抗性経路関連遺伝子として複数の候補遺伝子が見出された (表 I-4)。著者はこのうち、最も発現低下の顕著であった *Pik3r1* 遺伝子に注目した。この遺伝子はインスリン受容体の下流シグナルにおいて中心的役割を果たす PI3K をコードする遺伝子であり、インスリン抵抗性への関与が過去に報告されている [15,16]。そこで著者は、クエチアピンによるインスリン抵抗性は PI3K の阻害を介しており、ビタミン D は *Pik3r1* の発現低下に拮抗的に働いているのではないかと予想した。

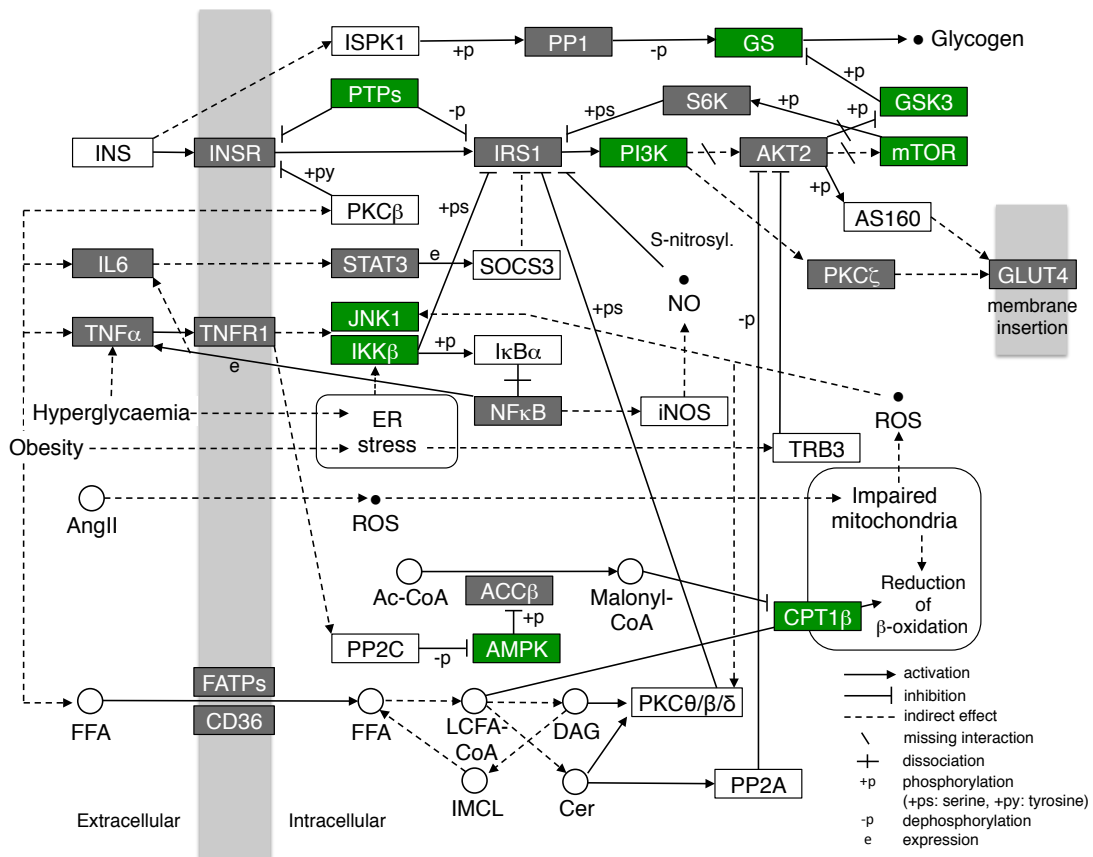


図 I-2. GEO 及び KEGG PATHWAY データベースの統合解析によるクエチアピン誘発インスリン抵抗性の原因遺伝子の予測

クエチアピン投与後のマイクロアレイデータを GEO データベースから入手し、その解析結果を KEGG PATHWAY データベースのインスリン抵抗性経路へマッピングした。クエチアピン投与により発現が変動しうると予想された遺伝子(表 I-4)を緑色で、有意な発現変動の認められなかった遺伝子を灰色で示している。図中の略称は以下の通りである。Ac-CoA = acetyl-coenzyme A. ACCβ = acetyl-coenzyme A carboxylase β. AMPK = AMP-activated protein kinase. AngII = angiotensin II. AS160 = Akt substrate of 160 kDa. Cer = ceramide. CPT1β = carnitine palmitoyltransferase 1 β. DAG = diacylglycerol. FATPs = fatty acid transport proteins. FFA = free fatty acid. GLUT4 = glucose transporter type 4. GS = glycogen synthase. GSK3 = glycogen synthase kinase 3. IKKβ = inhibitor of nuclear factor-kappa-B kinase subunit β. IL6 = interleukin 6. IMCL = intramyocellular lipid. iNOS = inducible nitric oxide synthase. INS = insulin. INSR = insulin receptor. IRS1 = insulin receptor substrate 1. ISPK1 = insulin-stimulated protein kinase 1. IκBα = nuclear factor-kappa-B inhibitor α. JNK1 = c-Jun N-terminal kinase 1. LCFA-CoA = long-chain fatty acyl-coenzyme A. mTOR = mechanistic target of rapamycin. NFκB = nuclear factor-κ-B. NO = nitric oxide. PI3K = phosphatidylinositol 3-kinase. PKC = protein kinase C. PP = protein phosphatase. PTPs = protein tyrosine phosphatases. ROS = reactive oxygen species. S6K = p70 ribosomal S6 kinase. SOCS3 = suppressor of cytokine signalling 3. STAT3 = signal transducer and activator of transcription 3. TNFR1 = tumor necrosis factor receptor 1. TNFα = tumor necrosis factor α. TRB3 = tribbles homolog 3.

Probe name	Gene symbol (Gene title)	Encoding protein	Relative expression	P value
NM_013005_P ROBE1	<i>Pik3r1</i> (phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 1)	PI3K	0.39 ± 0.13	0.0143
NM_013200_P ROBE1	<i>Cpt1b</i> (carnitine palmitoyltransferase 1b, muscle)	CPT1β	0.46 ± 0.01	0.0186
L27112_PROB E1	<i>Mapk9</i> (mitogen-activated protein kinase 9)	JNK1	0.46 ± 0.08	0.0021
X95577_PROB E1	<i>Prkab1</i> (protein kinase, AMP-activated, beta 1 non-catalytic subunit)	AMPK	0.50 ± 0.12	0.0176
X73653_PROB E1	<i>Gsk3b</i> (glycogen synthase kinase 3 beta)	GSK3	0.63 ± 0.05	0.0025
L37085_PROB E1	<i>Frap1</i> (FK506 binding protein 12-rapamycin associated protein 1)	mTOR	0.66 ± 0.03	0.0001
NM_013089_P ROBE1	<i>Gys2</i> (glycogen synthase 2)	GS	0.67 ± 0.06	0.0050
AI172465_PRO BE1	<i>Ptpn11</i> (protein tyrosine phosphatase, non-receptor type 11)	PTPs	0.70 ± 0.07	0.0173
AF115282_PR OBE1	<i>Ikkkb</i> (inhibitor of kappaB kinase beta)	IKKβ	0.70 ± 0.08	0.0396
Z29486_PROB E1	<i>Prkaa2</i> (protein kinase, AMP-activated, alpha 2 catalytic subunit)	AMPK	0.74 ± 0.05	0.0113
AI598719_PRO BE1	<i>Mapk8</i> (mitogen-activated protein kinase 8)	JNK1	1.30 ± 0.09	0.0489

表 I-4. インスリン抵抗性関連遺伝子の発現量に対するクエチアピン投与の影響を示す GEO マイクロアレイ解析データ

図 I-2 において統計学的に有意($P < 0.05$)な発現変動の認められた各遺伝子の詳細。データは対照群に対するクエチアピン投与群の相対的発現量を平均値±標準誤差で示している。N = 20–23 (対照群), N = 3 (クエチアピン投与群)。

グルコース負荷試験後のインスリン受容体下流遺伝子の発現に対する クエチアピン・ビタミン D 投与の影響

続いて著者は、クエチアピンによるインスリン抵抗性がビタミン D の併用により抑制されるという仮説についても薬理学的実験により検証した。マウスのグルコース負荷試験終了直後、インスリン依存性グルコース取込の主要部位である[17]骨格筋を抽出し、インスリン受容体下流シグナル経路遺伝子の mRNA 発現量を qRT-PCR 法により測定した(図 I-3)。統計学的に有意な差は認められなかったものの(Welch's *t*-test: $P=0.079$)、クエチアピン(10 mg/kg)投与マウスの骨格筋では *Pik3r1* mRNA の発現低下傾向が認められた(図 I-3c)。また、コレカルシフェロール添加餌(1200 IU vitamin D₃/day)飼育クエチアピン投与群ではその発現低下に対する有意な抑制が認められた(図 I-3c)。加えて、コレカルシフェロール添加餌飼育マウスではインスリン受容体をコードする *Insr* 遺伝子の有意な発現増加も認められ、既報[18]に一致する結果が確認された(図 I-3a)。*Irs1*, *Akt2* 及びグルコース輸送体をコードする *Slc2a4* といったその他のインスリン受容体下流シグナル遺伝子の発現に有意な変化は認められなかった(図 I-3b,d,e)。これらの実験結果から、ビタミン D の併用はクエチアピンによるインスリン抵抗性に対し *Pik3r1* の発現低下抑制を介して拮抗的に働くことが示唆され、GEO 及び KEGG PATHWAY データベース解析による予測に一致する結果が実験的にも確認された。

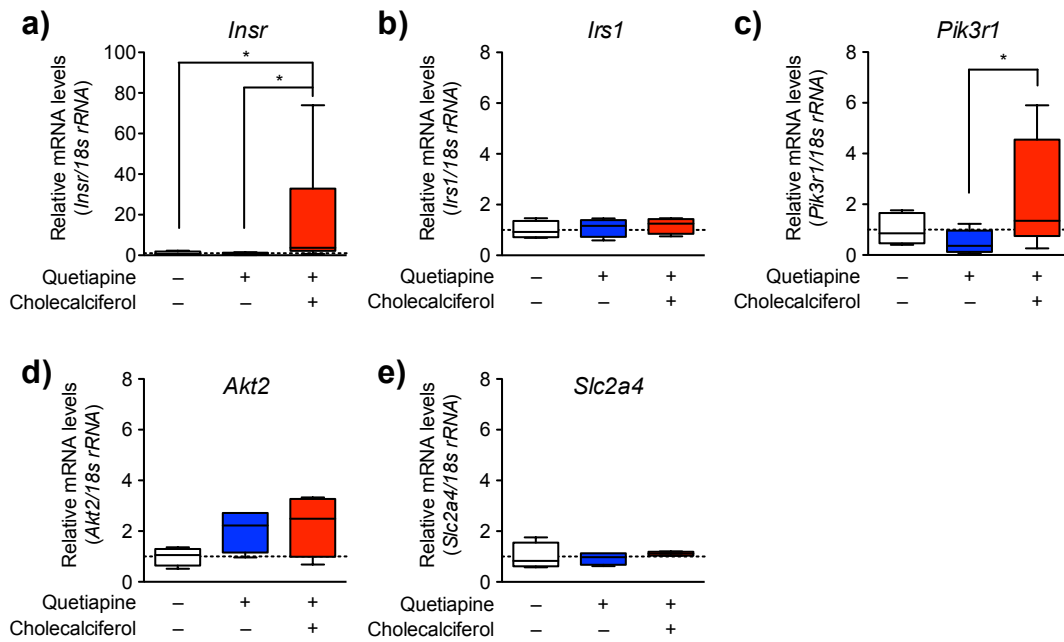


図 I-3. インスリン受容体シグナル経路遺伝子の発現に対するクエチアピン・ビタミン D 投与の影響

図 I-1 と同じく、ICR マウスを普通餌又はコレカルシフェロール添加餌(1200 IU vitamin D₃/day)で飼育しグルコース負荷試験を実施した。150 分間の試験終了直後にマウス骨格筋を回収し、(a) *Insr*, (b) *Irs1*, (c) *Pik3r1*, (d) *Akt2* 及び (e) *Slc2a4* mRNA の発現量を qRT-PCR 法により定量した。内部標準遺伝子は *18s rRNA* とし、対照群に対する相対的発現量を各遺伝子について算出した。データは箱ひげ図(上から順に最大値, 第 3 四分点 (75%), 中央値, 第 1 四分点(25%), 最小値)で示している。N = 4-8, * P < 0.05.

グルコース取込試験におけるクエチアピン・ビタミン D 処置の影響と PI3K シグナリング経路の関与の検討

最後に、著者は C2C12 培養骨格筋細胞を用いたグルコース取込試験により仮説の更なる補強を試みた(図 I-4). クエチアピン(1-100 μM)の1時間処置により有意かつ濃度依存的なインスリン誘発グルコース取込の抑制が認められたことから、クエチアピンによりインスリン抵抗性が引き起こされていることが示唆された(図 I-4a). このクエチアピン(100 μM)によるインスリン抵抗性は生体内活性型ビタミン D であるカルシトリオール(0.1-10 nM)の24時間前処置により有意かつ濃度依存的に改善された(図 I-4b). カルシトリオールの前処置自体によるインスリン誘発グルコース取込への影響は認められなかった(図 I-4c). 加えて、カルシトリオール併用によるインスリン抵抗性の改善は、AMPK 阻害薬である compound C (20 μM)存在下では変化が認められなかった一方、PI3K 阻害薬である LY294002 (20 μM)存在下では有意に抑制された(図 I-4d). これらの結果より、クエチアピン誘発インスリン抵抗性に対するビタミン D 併用の改善作用は PI3K 経路の活性化を介していることが細胞レベルで確認された.

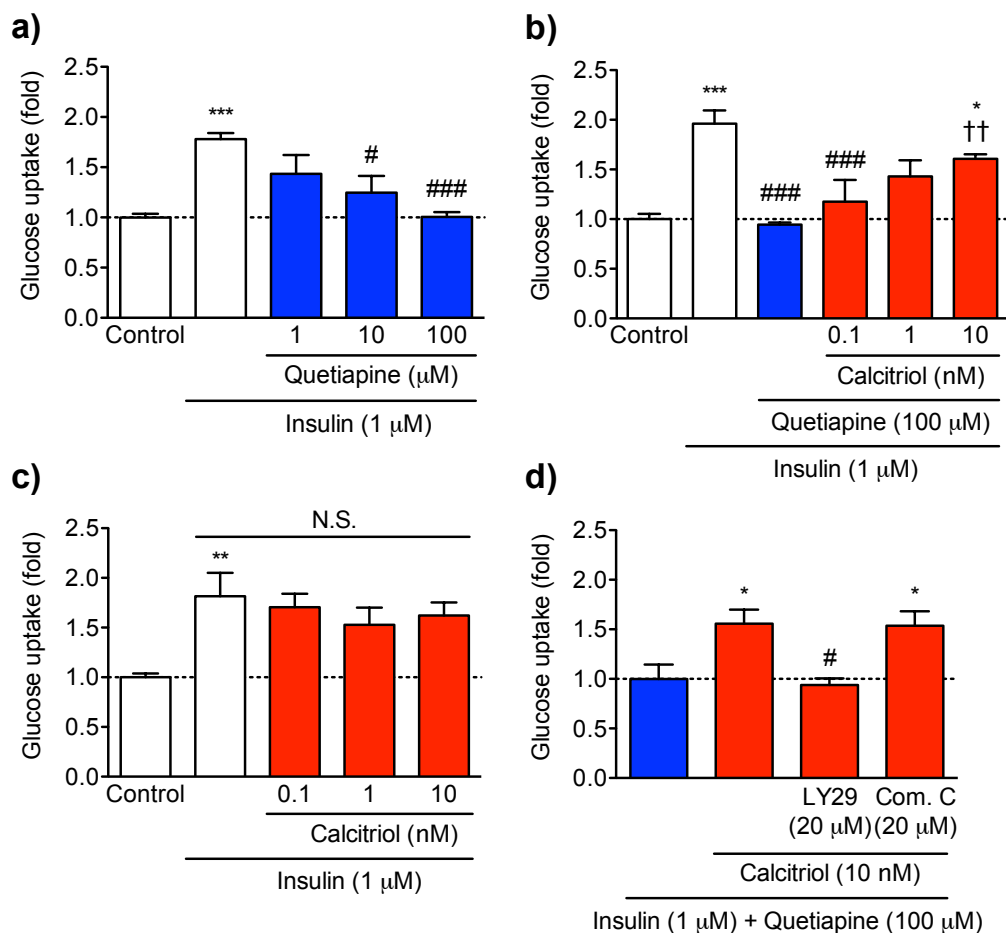


図 1-4. C2C12 筋管細胞におけるクエチアピン誘発インスリン抵抗性に対するカルシトリオール処置の影響及び PI3K シグナリング経路の関与の検討

グルコース取込量は 2-NBDG を用いて評価した. (a) インスリン誘発グルコース取込に対するクエチアピン処置の影響. 分化 C2C12 細胞に対しクエチアピン(1–100 μM)を 1 時間処置した後, インスリン(1 μM)で 15 分間刺激した. $N=10$. *** $P<0.001$ vs. 対照群; # $P<0.05$, ### $P<0.001$ vs. インスリン処置群. (b) クエチアピン誘発インスリン抵抗性に対するカルシトリオール処置の影響. 分化 C2C12 細胞に対しカルシトリオール(0.1–10 nM)を 24 時間前処置し, クエチアピン(100 μM)を 1 時間処置した後, インスリン(1 μM)で 15 分間刺激した. $N=10$. * $P<0.05$, *** $P<0.001$ vs. 対照群; ### $P<0.001$ vs. インスリン処置群; †† $P<0.01$ vs. クエチアピン処置群. (c) インスリンによるグルコース取込に対するカルシトリオール処置の影響. 分化 C2C12 細胞に対しカルシトリオール(0.1–10 nM)を 24 時間処置した後, インスリン(1 μM)で 15 分間刺激した. $N=9$. ** $P<0.01$ vs. 対照群; N.S. = not significant. (d) カルシトリオールによるクエチアピン誘発インスリン抵抗性の改善に対する LY294002 (PI3K 阻害薬)及び compound C (AMPK 阻害薬)の影響. 分化 C2C12 細胞にカルシトリオール(10 nM)を 24 時間前処置し, LY294002 (LY29; 20 μM) 又は compound C (Com. C; 20 μM)と共にクエチアピン(100 μM)を 1 時間処置した後, インスリン(1 μM)で 15 分間刺激した. $N=9-13$, * $P<0.05$ vs. クエチアピン処置群; # $P<0.05$ vs. クエチアピン・カルシトリオール併用群. データは対照群(a–c)又はインスリン及びクエチアピン処置群(d)に対する相対値として平均値と標準誤差で示している.

考察

有害事象ビッグデータ解析を分子薬理学実験に応用した本研究は、「ビタミン D が抗精神病薬誘発高血糖に対し予防作用を有しており、その機序として *Pik3r1* 遺伝子の発現上昇を介したインスリン抵抗性の改善が考えられる」ということを初めて示唆するものである。

非定型抗精神病薬による代謝障害に対するビタミン D の効果についてはほとんど知られていない。抗精神病薬誘発代謝障害に対するビタミン D の有効性を評価した臨床研究は一部存在するが、どちらもサンプルサイズが不十分であるため統計学的な推論には限界がある[19,20]。また、抗精神病薬による高血糖に対するビタミン D 欠乏症の関与を示唆する研究も存在する[21]が、この研究は動物モデルのみに基づき結論を下しており、「クロザピン投与により体重増加が抑制される」といった、臨床の知見とは一致しない結果も得られてしまっている。これらの研究に対し、本研究は臨床データベース解析と薬理学的な *in vivo/in vitro* 実験の両方を組み合わせているため、抗精神病薬誘発高血糖に対するビタミン D 併用の予防効果について分子レベルから臨床レベルまでのエビデンスを含んでいるという特徴がある。また、本研究成果はビタミン D が 2 型糖尿病の進行を予防しうることを示唆する先行研究[22]とも整合性がとれるものである。

北米肥満学会による報告[1]では、クロザピンとオランザピンについては糖尿病リスクが高く、クエチアピンとリスペリドンについては一致した見解は得られず、アリピプラゾールとジプラシドンについては糖尿病リスクの増加は認められないとされている。しかしながら、本研究の FAERS 解析においては、クロザピンの使用と糖尿病報告増加との間の相関は比較的弱い結果となった(表 I-1)。クロザピン使用症例の糖尿病報告が比較的少なかった原因としては、クロザピンの使用には血液検査が必須である[23]ことから、クロザピン使用患者における耐糖能異常は早期発見され、重篤化に至らなかったからという説明が可能かもしれない。一方で、アリピプラゾールとジプラシドンについては糖尿病報告増加との間に比較的強い相関が認められたが、これは抗精神病薬の多剤併用に起因すると考えられる。実際、FAERS の追加解析を実施したところ、ジプラシドンを使用し糖尿病が報告されている症例の約 85% (1368/1613)にはクエチアピンも処方されており、糖尿病の主被疑薬としては 77.5% (1250/1613)がクエチアピンを指定していた一方、ジプラシドンが主被疑薬として指定されていたのは 5.6% (91/1613)のみであった。

ビタミンDは核内のビタミンD受容体と結合することで転写因子として機能し標的遺伝子の発現を制御する[24]. 本研究において、著者は抗精神病薬誘発耐糖能異常及びインスリン抵抗性に関する *in vivo/in vitro* の薬理的病態モデルを作製し、ビタミンD投与による *Pik3r1* 遺伝子の発現上昇及びPI3Kの機能的関与を示した. 骨格筋における *Pik3r1* 及び *Pik3r2* 両遺伝子を欠損させたマウスはPI3Kシグナリング異常及びインスリン抵抗性を示すことが報告されている[25]. また、AktはPI3Kの下流で機能しGLUT4の膜移行を促進することで細胞内へのグルコース取込を増加させる[26]. これらの知見をまとめると、*Pik3r1* 遺伝子の発現低下はGLUT4を介したグルコース取込異常を引き起こすことで高血糖症状を増悪させると考えられる. 加えて、ビタミンD投与による *Insr* 遺伝子の発現亢進も抗精神病薬誘発耐糖能異常の改善に寄与している可能性が考えられる.

しかしながら、本研究は抗精神病薬及びビタミンDの作用機序における *Pik3r1* 遺伝子以外の関与を否定するものではない. 例えば、非定型抗精神病薬・ビタミンDのいずれも、インスリンシグナル経路タンパク質(IRS1やAkt等)のリン酸化制御によりノンジェノミックに作用しうることが報告されている[27-29]. 加えて、抗精神病薬が膵臓β細胞のアポトーシスを誘発しインスリン分泌を抑制することや[30,31]、ビタミンDが膵臓β細胞の機能を改善することも報告されている[22]. それ故、インスリン抵抗性と *Pik3r1* 遺伝子との関与については更なる調査が必要であると考えられる.

FAERSやGEOといったデータベースは医薬品の毒性研究に有用なツールではあるものの、データマイニングによる仮説生成には様々な制約が存在する. 例えば、FAERSの解析では適応症バイアスの影響により、クエチアピン誘発糖尿病に対するメトホルミンの有効性を検出することはできなかった(OR = 5.29, 95% CI = 4.90-5.71). また、GEOには大量のトキシコゲノミクスデータセットが登録されているものの、実験プロトコルの多様性を考慮すれば依然として使用可能なデータには限界がある. しかしながら、これらの欠点についてはシステムズ薬理学データベースの統合解析[32]やトキシコゲノミクスデータのメタ・アナリシス[33]といった手法の改善により克服しうると考えられる. データマイニングの活用は抗精神病薬誘発高血糖のように、従来型の臨床研究や動物実験による検証が困難な複雑な病態における対処法の探索・分子メカニズムの解明に有用なツールとなると考えられる.

まとめると、データマイニングによる予測及び *in vivo/in vitro* 実験による検証

を組み合わせた本研究により、抗精神病薬誘発高血糖及びそれに伴うインスリン抵抗性に対するビタミン D 併用療法の有効性が新たに示唆された。このように、有害事象ビッグデータ解析と分子薬理学的実験検証を組み合わせることで、薬物有害事象に対する実用的な対処法をその分子メカニズムと共に提唱することが可能となると考えられる。

第二章

分子薬理学研究ツールとしての 有害事象ビッグデータの解析手法開発

序論

第一章で提示した有害事象ビッグデータの薬剤疫学的解析は、非定型抗精神病薬使用患者における糖尿病のような医療ニーズの高い薬物有害事象に対する新たな解決策を提示するのみならず、ビタミン D のような既存薬の新たな薬効探索(ドラッグ・リポジショニング)や、*Pik3r1* 遺伝子のような病態形成に關与する新たな分子機序の解明にも応用できる有用なツールとなると考えられる。しかしながら、薬物有害事象研究をこれまで中心的に担ってきた薬剤疫学分野は多種多様な背景を有する患者データの後方視的解析を中心に発展してきたため、統制された条件下で動物・細胞実験を実施し処置群と対照群の比較データを取得する分子薬理学分野に比べ、扱うデータの性質、表現、解析及び解釈に必要な「常識」が大きく異なる。そこで本章では、薬剤疫学分野におけるこれらの「常識」を分子薬理学研究者に伝わる形に整理し、必要な統計解析技術、適切な解釈の仕方、手法の限界及びその対策について検討を行った。

方法

FAERS データベースの解析

基本的には第一章と同様であるであるが、第一章での研究から以下のように修正・更新されたため第一章と数値が異なる。

1. FAERS のデータ保管場所が <https://fis.fda.gov/extensions/FPD-QDE-FAERS/FPD-QDE-FAERS.html> に変更され, FDA によりデータが若干修正された。
2. 2014 年第 3 四半期(14Q3)から 2018 年第 4 四半期(18Q4)までのデータを新たに追加し, 合計 11,904,706 件の報告から重複症例を取り除いた 9,948,368 件の症例について解析した。
3. ライフサイエンス辞書シソーラスによる薬物名の統一ファイルを新たに更新した。

JADER データベースの解析

JADER データセットは PMDA のウェブサイト(<https://www.info.pmda.go.jp/fukusayoudb/CsvDownload.jsp>)からダウンロードした。2004 年第 1 四半期から 2018 年第 3 四半期までに JADER データベースに登録された計 561,122 件の報告を解析に用いた。

統計解析

FAERS 及び JADER データベースの統計解析は R version 3.4.3 Software (R foundation for Statistical Computing)及び Prism 8 Software (GraphPad)を用いた。Prism は分子薬理学分野で広く用いられている統計ソフトであり, 分子薬理学分野で用いられる解析手法と薬剤疫学分野で用いられる解析手法との対応関係を明示するため, 図の一部には Prism 様のテーブル表現を用いた。R を用いた群間比較には stats パッケージの `t.test`, `wilcox.test`, `chisq.test`, `fisher.test`, `glm`, `summary`, `anova` 関数を用いた。 P 値の計算は必要に応じて stats パッケージの `pnorm`, `pchisq`, `pt`, `pf`, `phyper` 関数を用いた。傾向スコアの計算には `speedglm` パッケージ[36]の `shglm` 及び `speedglm` 関数によるロジスティック回帰分析を, マッチングには `Matching` パッケージ[37]の `Matching` 関数及び `MatchIt` パッケージ[38]の `matchit` 関数による最近傍法を用い, `caliper` 変数には 0.25 を指定した。その他, 必要に応じて自作関数を用いて解析を行った。

結果

被疑薬・有害事象の 2 項関係の表現方法

最初に、被疑薬(クエチアピン)と有害事象(糖尿病)との間の 2 項関係の表現方法について検討する(図 II-1). FAERS のような有害事象自発報告データベースを用いた研究では一般に、興味のある薬物の有無・有害事象の有無で報告を分類し、図 II-1a のような 2×2 分割表(contingency table)を用いてデータを表現する[34]. 一方で、分子薬理学分野における動物・細胞実験では一般に、データを平均値±標準誤差及び N に要約しグラフに可視化することが多く、分割表を用いた解析を行う機会は比較的稀である. 数学的には、図 II-1a のような分割表は糖尿病ありを 1, 糖尿病無しを 0 としたダミー変数を用いることにより、分子薬理学実験で一般に用いられるカラム型テーブルに変換できる(図 II-1b). このような対応関係を理解しておけば図 II-1a のような分割表データを分子薬理学実験と同じく平均値±標準誤差及び N に要約しグラフに可視化することができる(図 II-1c,d).

図 II-1c の図から、FAERS 解析で扱うデータは分子薬理学研究で扱う動物・細胞実験データに比べ以下の点で大きく異なることが分かる.

- (1) 測定値が 0 と 1 しか取らない
- (2) N が極めて大きい
- (3) 各群の N が不均等である

上記の 3 点に加え、図 II-3 で後述するように

- (4) 各群の背景因子が一致しない

という相違点も存在する. これらの相違点(1)及び(2)については図 II-2 で、(3)及び(4)については図 II-3 で検討する.

以降の検討では、分子薬理学実験に合わせ各群を平均値±標準誤差及び N に要約して表現する. ただし平均値は割合なので 100 倍してパーセント表記にした方が一般的には理解されやすいと考えられる. また、エラーバーとしては±標準誤差よりも統計学的に優れた表現方法が存在する(図 II-1e)[39,40].

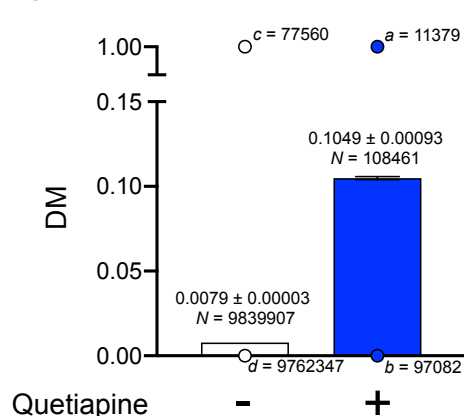
a) Contingency table expression

Table format: Contingency		Outcome A	Outcome B
		DM	No DM
		Y	Y
1	Quetiapine	11379	97082
2	No quetiapine	77560	9762347

b) Column table expression

Table format: Column		Group A	Group B
		No quetiapine	Quetiapine
		Y	Y
1	Title	1	1
2	Title	1	1
3	Title	1	1
11378	Title	1	1
11379	Title	1	1
11380	Title	1	0
11381	Title	1	0
77559	Title	1	0
77560	Title	1	0
77561	Title	0	0
77562	Title	0	0
108460	Title	0	0
108461	Title	0	0
108462	Title	0	0
9839906	Title	0	0
9839907	Title	0	0

c) Bar graph visualization



d) Mean, SEM, N summarization

Table format: Grouped		Group A			Group B		
		No quetiapine			Quetiapine		
	Title	Mean	SEM	N	Mean	SEM	N
1	Title	0.0079	0.00003	9839907	0.1049	0.00093	108461

e) Statistically better error bars

Fraction of total		A		
		DM		
	Title	Mean	Upper Limit	Lower Limit
1	No quetiapine	0.0079	0.0079	0.0078
2	Quetiapine	0.1049	0.1068	0.1031

図 II-1. クエチアピン・糖尿病の 2 項関係の表現方法

(a) クエチアピン使用と糖尿病報告との間の関係を示す 2×2 分割表. (b) 糖尿病有りを 1, 糖尿病無しを 0 として(a)の 2×2 分割表をカラム型テーブル形式に変換したもの. (c) カラム型テーブル(b)の棒グラフによる可視化. バーは平均値±標準誤差を示している. (d) カラム型テーブル(b)の平均値, 標準誤差及び N による要約. (e) Prism 8 による 95% Score 信頼区間計算結果. 図 II-7e-h に示すように, 併用薬探索の場合には 0 セルを含む併用薬が候補として重要となってくるが, この場合標準誤差も 0 となるため, 平均値±標準誤差や平均値±1.96×標準誤差(95% Wald 信頼区間)による表現方法ではエラーバーも 0 となってしまう. このような場合でもエラーバーを計算する方法として, 統計ソフト Prism 8 には Wilson 法等による信頼区間計算機能が実装されており, 統計ソフト R でも標準関数 prop.test 等により計算可能である.

被疑薬・有害事象の 2 項関係における群間の比較

続いて、被疑薬(クエチアピン)使用群と非使用群における有害事象(糖尿病)報告割合の比較方法について検討する(図 II-2). 図 II-1c のような 2 群間を比較する場合、分子薬理学実験では Student の t 検定を用いるのが一般的である(図 II-2a). 他に Welch の t 検定や Mann-Whitney 検定等が用いられることもあり、各検定法による P 値を計算すると図 II-2b のようになる. ただし今回扱うデータは図 II-1 の相違点(1)で述べたように測定値が 0 と 1 の値しか取らない 2 値データ(binary data)であるため、 t 検定でなく Pearson の χ^2 検定(図 II-2c)や Fisher の直接確率検定等が用いられ、それらの P 値を計算すると図 II-2d のようになる. また現代では一般化線形モデル(GLM)を用いた解析手法が広く用いられており、差のみならず比・オッズ比を指標に計算することや、Wald 法、Score 法、尤度比検定(LRT)法による P 値計算が可能である(図 II-2e).

これらの解析手法の数学的な性質の違いについては統計学の専門書[39,40]に譲るとして、ここでは薬剤疫学・分子薬理学の慣例の違いについて整理する.

- (5) 有害事象データベースの薬剤疫学的解析では比(Proportional Reporting Ratio, PRR)やオッズ比(Reporting Odds Ratio, ROR)を指標とするのが一般的である[34]のに対し、分子薬理学分野で広く用いられる Student の t 検定は差を指標とする検定法である.
- (6) 有害事象データベースの薬剤疫学的解析におけるオッズ比の信頼区間計算は一般に Wald 法が用いられる(図 II-2e 橙色部)[8,39,40]のに対し、分子薬理学研究で一般的な Student の t 検定に最も近い手法は Score 検定である(図 II-2 黄色部).
- (7) 疫学分野では一般に効果量(2 群間の差・比・オッズ比等)の 95%信頼区間を強調し P 値を記載しないことも多い[41]のに対し、分子薬理学分野では P 値をアスタリスク(*)に要約してグラフ上に記入するのが一般的である.

これらの相違点が組み合わさった結果、有害事象データベースの薬剤疫学的解析において一般的なオッズ比と Wald 信頼区間を用いた群間比較表現(表 I-1)は、Student の t 検定による P 値をグラフ上に記入し 2 群間の差の有意性を表現する分子薬理学分野の慣例からすると大きく異なる印象を与えることになる. これらの薬剤疫学・分子薬理学の慣例の相違点(5-7)については図 II-7 において再検討する.

上述の相違点に加えて、図 II-1 の相違点(2)で指摘したように N が極めて大きい
ため、

- (8) 分子薬理学実験のように N が小さい($N = 10$ 程度)場合、 $P < 0.001$ を伝統的に“extremely significant”等と表現するが、有害事象ビッグデータ解析で扱うデータは $N = 10000000$ 程度となるため、 0.001 のような伝統的な閾値を下回ったとしても P 値が十分に小さいことを証明しているとは限らない。

という P 値の解釈の注意点も関与してくることになる。このため、有害事象ビッグデータ解析において統計学的有意性を示すには 0.05 や 0.01 , 0.001 といった伝統的な閾値との比較よりも、他の薬物や有害事象、ポジティブ/ネガティブコントロールデータと比較し P 値がどの程度小さいのかを推論する過程が重要となり、そのためには P 値をアスタリスク(*)に要約せず図 II-2b,d,e のように直接示す方が比較の際に有用となる。本研究の場合、全薬物中糖尿病報告と最も統計学的に有意な相関の認められる薬物がクエチアピンであり (Score test $P = 10^{-24757}$, 全 2976 薬物中最小)、逆に全有害事象中クエチアピン使用と最も統計学的に有意な相関の認められる有害事象が糖尿病である (Score test $P = 10^{-25108}$, 全 20693 有害事象中最小)。先行研究[8]におけるロシグリタゾンと心筋梗塞も同様の関係にある。

a) Student's *t* test

Unpaired <i>t</i> test	
Unpaired <i>t</i> test	
P value	<0.0001
P value summary	****
Significantly different ($P < 0.05$)?	Yes
One- or two-tailed P value?	Two-tailed
<i>t</i> , df	$t=339.6$, $df=9948366$
How big is the difference?	
Mean of Column A	0.007882
Mean of Column B	0.1049
Difference between means (B - A) \pm SEM	0.09703 ± 0.0002857
95% confidence interval	0.09647 to 0.09759

c) Pearson's χ^2 test

Contingency	
P value and statistical significance	
Test	Chi-square
Chi-square, df	111397, 1
<i>z</i>	337.6
P value	<0.0001
P value summary	****
One- or two-sided	Two-sided
Statistically significant ($P < 0.05$)?	Yes

b) Column analyses

Student's *t* test P value = 10^{-24900}
 Welch's *t* test P value = 10^{-2251}
 Mann-Whitney test approximate P value = 10^{-24757}
 Mann-Whitney test exact P value = 10^{-8159}

d) Contingency table analyses

Fisher's exact test P value = 10^{-8159}
 Yates' χ^2 test P value = 10^{-24754}
 Pearson's χ^2 test P value = 10^{-24757}

e) Generalized linear models (GLMs) for binary data

		Difference +0.097	Ratio $\times 13.31$	Odds ratio $\times 14.75$
Wald test	P value	10^{-2361}	10^{-15912}	10^{-14151}
	95% CI	0.095 to 0.099	13.06 to 13.56	14.45 to 15.06
Score test	P value	10^{-24757}	10^{-24757}	10^{-24757}
	95% CI			
Likelihood ratio test (LRT)	P value	10^{-8159}	10^{-8159}	10^{-8159}
	95% CI	0.095 to 0.099	13.06 to 13.56	14.45 to 15.06

図 II-2. クエチアピン・糖尿病の 2 項関係における群間の比較

(a) 図 II-1c データに Student の *t* 検定を適用した場合の解析結果. (b) 分子薬理学実験で一般に用いられる各検定手法の P 値計算結果. (c) 図 II-1a 分割表に Pearson の χ^2 検定を適用した場合の解析結果. (d) 分割表の解析で一般に用いられる各検定手法の P 値計算結果. (e) 一般化線形モデルを用いた場合の効果量, P 値及び 95%信頼区間計算結果. Score 信頼区間は解析ソフトウェアを用いても一般に計算困難であり表には記載していない. (a)及び(c)は Prism 8, (b), (d)及び(e)は R による解析結果.

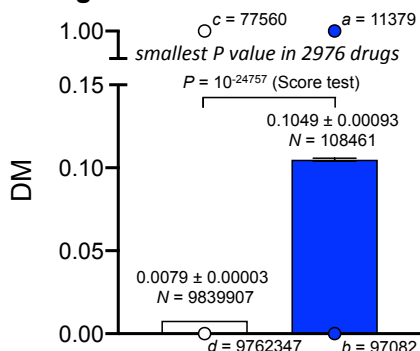
被疑薬・有害事象の2項関係における背景の調整

続いて、被疑薬(クエチアピン)使用群と非使用群の背景の調整について検討する(図 II-3). 分子薬理学分野では基本的に、各群の N と条件が均等となるように実験系を構築する. しかしながら図 II-1 の相違点(3)で確認した通り、FAERS データベース内の症例を単純にクエチアピン使用の有無で分割した場合、各群の N のバランスは著しく不均等となる. 加えて、年齢、性別、原疾患、併用薬といった交絡となりうる背景因子が両群で一致しないという相違点(4)も存在する. 例えば図 II-3a の場合、クエチアピン非使用群に比べクエチアピン使用群の方が統合失調症の罹患割合が高く(54137/9839907 (0.55%) vs. 6642/108461 (6.12%), Score test $P = 10^{-11920}$), オランザピンの使用割合も高い(51792/9839907 (0.53%) vs. 8067/108461 (7.44%), Score test $P = 10^{-18608}$). これは、クエチアピン・オランザピンが共に統合失調症の治療薬であることを考えれば当然の結果であるが、統合失調症の罹患[42]やオランザピンの使用についても糖尿病との関与が報告されているため、クエチアピン使用群における糖尿病報告割合の増加が、クエチアピンの使用に起因しているのか、それとも様々な背景因子の違いに起因しているのかを区別することができない.

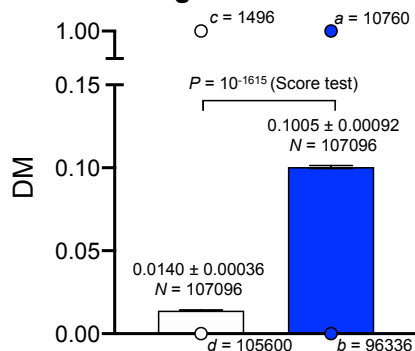
このような状況に対し、薬剤疫学分野では傾向スコアマッチング(propensity score matching)のような統計学的手法を適用することで、各群の背景を技術的に調整する. 図 II-3b は、上述の統合失調症のような2群間で罹患割合の異なる疾患上位200件(Score 検定 P 値順)と、オランザピンのような2群間で使用割合の異なる薬物上位200件(Score 検定 P 値順)の計400種類の情報を利用し、両群の N と背景因子が均等になるようにマッチングを実施したものである. このように、背景因子として疾患と薬物を統一すれば、年齢や性別等の他の背景因子についても自動的に調整されることが先行研究[7]で示されている. マッチング実施後においてもクエチアピン使用群の方が非使用群に比べ糖尿病報告割合が高いことから、クエチアピンの使用と糖尿病報告割合増加との間の相関が改めて確認できる.

傾向スコアマッチングにより得られる結果は分子薬理学研究者にとっても理解しやすく、一見理想的な方法のように見えるが、実際には数多くの限界が存在する. 調整の仕方や解析ソフトにより結果が大きく左右されるため、調整無し(crude)で候補をある程度絞り込み、調整しても(adjusted)相関が認められることを確認する、という多段階の解析を実施するのが現実的な対処となる[43].

a) Before matching



b) After matching



	Quetiapine -		Quetiapine +		Quetiapine -		Quetiapine +	
Mean age (years old)	54.61	48.02	51.58	48.09	51.58	48.09	51.58	48.09
Female (%)	55.41	55.88	54.32	55.89	54.32	55.89	54.32	55.89
Bipolar disorder N (%)	42251 (0.43)	14541 (13.41)	10360 (9.67)	13789 (12.88)	10360 (9.67)	13789 (12.88)	10360 (9.67)	13789 (12.88)
Depression N (%)	180940 (1.84)	15343 (14.15)	15872 (14.82)	14669 (13.70)	15872 (14.82)	14669 (13.70)	15872 (14.82)	14669 (13.70)
Sleep disorder N (%)	28879 (0.29)	5960 (5.50)	5120 (4.78)	5495 (5.13)	5120 (4.78)	5495 (5.13)	5120 (4.78)	5495 (5.13)
Schizophrenia N (%)	54137 (0.55)	6642 (6.12)	6356 (5.94)	6306 (5.89)	6356 (5.94)	6306 (5.89)	6356 (5.94)	6306 (5.89)
Psychotic disorder N (%)	13308 (0.14)	3160 (2.91)	2563 (2.39)	2972 (2.78)	2563 (2.39)	2972 (2.78)	2563 (2.39)	2972 (2.78)
Bipolar I disorder N (%)	5445 (0.06)	2082 (1.92)	1524 (1.42)	1945 (1.81)	1524 (1.42)	1945 (1.81)	1524 (1.42)	1945 (1.81)
Anxiety N (%)	88897 (0.90)	7557 (6.97)	7245 (6.77)	7082 (6.61)	7245 (6.77)	7082 (6.61)	7245 (6.77)	7082 (6.61)
Insomnia N (%)	63188 (0.64)	5171 (4.77)	5159 (4.82)	4914 (4.59)	5159 (4.82)	4914 (4.59)	5159 (4.82)	4914 (4.59)
Affective disorder N (%)	7735 (0.08)	1687 (1.56)	1292 (1.21)	1506 (1.41)	1292 (1.21)	1506 (1.41)	1292 (1.21)	1506 (1.41)
Agitation N (%)	5391 (0.05)	1364 (1.26)	1040 (0.97)	1235 (1.15)	1040 (0.97)	1235 (1.15)	1040 (0.97)	1235 (1.15)
Schizoaffective disorder N (%)	8568 (0.09)	1645 (1.52)	1426 (1.33)	1571 (1.47)	1426 (1.33)	1571 (1.47)	1426 (1.33)	1571 (1.47)
Mental disorder N (%)	3562 (0.04)	1064 (0.98)	823 (0.77)	973 (0.91)	823 (0.77)	973 (0.91)	823 (0.77)	973 (0.91)
Mania N (%)	2281 (0.02)	839 (0.77)	630 (0.59)	781 (0.73)	630 (0.59)	781 (0.73)	630 (0.59)	781 (0.73)
Post-traumatic stress disorder N (%)	4351 (0.05)	1097 (1.01)	897 (0.84)	1019 (0.95)	897 (0.84)	1019 (0.95)	897 (0.84)	1019 (0.95)
Hallucination N (%)	6507 (0.07)	1215 (1.12)	1047 (0.98)	1166 (1.09)	1047 (0.98)	1166 (1.09)	1047 (0.98)	1166 (1.09)
Parkinson's disease N (%)	37321 (0.38)	2499 (2.30)	3110 (2.90)	2475 (2.31)	3110 (2.90)	2475 (2.31)	3110 (2.90)	2475 (2.31)
Bipolar II disorder N (%)	1339 (0.01)	454 (0.42)	329 (0.31)	426 (0.40)	329 (0.31)	426 (0.40)	329 (0.31)	426 (0.40)
Major depression N (%)	16262 (0.17)	1558 (1.44)	1775 (1.66)	1493 (1.40)	1775 (1.66)	1493 (1.40)	1775 (1.66)	1493 (1.40)
Delirium N (%)	1108 (0.01)	375 (0.35)	325 (0.30)	354 (0.33)	325 (0.30)	354 (0.33)	325 (0.30)	354 (0.33)
Paranoia N (%)	962 (0.01)	348 (0.32)	230 (0.21)	302 (0.28)	230 (0.21)	302 (0.28)	230 (0.21)	302 (0.28)
Mood swings N (%)	2246 (0.02)	512 (0.47)	372 (0.35)	461 (0.43)	372 (0.35)	461 (0.43)	372 (0.35)	461 (0.43)
Prophylaxis of neural tube defect N (%)	2770 (0.03)	566 (0.52)	428 (0.40)	546 (0.51)	428 (0.40)	546 (0.51)	428 (0.40)	546 (0.51)
Delusion N (%)	5563 (0.06)	792 (0.73)	781 (0.73)	776 (0.72)	781 (0.73)	776 (0.72)	781 (0.73)	776 (0.72)
Borderline personality disorder N (%)	785 (0.01)	307 (0.28)	211 (0.20)	297 (0.28)	211 (0.20)	297 (0.28)	211 (0.20)	297 (0.28)
Hallucination, auditory N (%)	789 (0.01)	307 (0.28)	181 (0.17)	265 (0.25)	181 (0.17)	265 (0.25)	181 (0.17)	265 (0.25)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Lithium bicarbonate N (%)	17609 (0.18)	5114 (4.72)	3902 (3.64)	4788 (4.47)	3902 (3.64)	4788 (4.47)	3902 (3.64)	4788 (4.47)
Balproic acid N (%)	73131 (0.74)	9918 (9.14)	8630 (8.06)	9324 (8.71)	8630 (8.06)	9324 (8.71)	8630 (8.06)	9324 (8.71)
Clonazepam N (%)	98588 (1.00)	11454 (10.56)	11078 (10.34)	10853 (10.13)	11078 (10.34)	10853 (10.13)	11078 (10.34)	10853 (10.13)
Olanzapine N (%)	51792 (0.53)	8067 (7.44)	6424 (6.00)	7486 (6.99)	6424 (6.00)	7486 (6.99)	6424 (6.00)	7486 (6.99)
Risperidone N (%)	80491 (0.82)	8932 (8.24)	7506 (7.01)	8280 (7.73)	7506 (7.01)	8280 (7.73)	7506 (7.01)	8280 (7.73)
Lithium N (%)	10640 (0.11)	3027 (2.79)	2454 (2.29)	2851 (2.67)	2454 (2.29)	2851 (2.67)	2454 (2.29)	2851 (2.67)
Haloperidol N (%)	23146 (0.24)	4380 (4.04)	3595 (3.36)	3962 (3.70)	3595 (3.36)	3962 (3.70)	3595 (3.36)	3962 (3.70)
Aripiprazole N (%)	65201 (0.66)	7370 (6.80)	6684 (6.24)	6863 (6.41)	6684 (6.24)	6863 (6.41)	6684 (6.24)	6863 (6.41)
Lorazepam N (%)	101031 (1.03)	8919 (8.22)	8954 (8.36)	8459 (7.90)	8954 (8.36)	8459 (7.90)	8954 (8.36)	8459 (7.90)
Lamotrigine N (%)	71375 (0.73)	7346 (6.77)	6885 (6.43)	7025 (6.56)	6885 (6.43)	7025 (6.56)	6885 (6.43)	7025 (6.56)
Trazodone N (%)	58099 (0.59)	6582 (6.07)	6074 (5.67)	6105 (5.70)	6074 (5.67)	6105 (5.70)	6074 (5.67)	6105 (5.70)
Citalopram N (%)	158261 (1.61)	11289 (10.41)	12119 (11.32)	10736 (10.02)	12119 (11.32)	10736 (10.02)	12119 (11.32)	10736 (10.02)
Venlafaxine N (%)	80983 (0.82)	7801 (7.19)	7594 (7.09)	7375 (6.89)	7594 (7.09)	7375 (6.89)	7594 (7.09)	7375 (6.89)
Ziprasidone N (%)	14090 (0.14)	3127 (2.88)	2076 (1.94)	2755 (2.57)	2076 (1.94)	2755 (2.57)	2076 (1.94)	2755 (2.57)
Mirtazapine N (%)	42553 (0.43)	5396 (4.98)	5153 (4.81)	5076 (4.74)	5153 (4.81)	5076 (4.74)	5153 (4.81)	5076 (4.74)
Sertraline N (%)	108948 (1.11)	7622 (7.03)	8071 (7.54)	7213 (6.74)	8071 (7.54)	7213 (6.74)	8071 (7.54)	7213 (6.74)
Benzatropine N (%)	6822 (0.07)	1807 (1.67)	1337 (1.25)	1649 (1.54)	1337 (1.25)	1649 (1.54)	1337 (1.25)	1649 (1.54)
Bupropion N (%)	85686 (0.87)	6217 (5.73)	6321 (5.90)	5810 (5.43)	6321 (5.90)	5810 (5.43)	6321 (5.90)	5810 (5.43)
Chlorpromazine N (%)	5706 (0.06)	1479 (1.36)	886 (0.83)	1257 (1.17)	886 (0.83)	1257 (1.17)	886 (0.83)	1257 (1.17)
Fluoxetine N (%)	73930 (0.75)	5476 (5.05)	5439 (5.08)	5108 (4.77)	5439 (5.08)	5108 (4.77)	5439 (5.08)	5108 (4.77)
Diazepam N (%)	68352 (0.69)	5224 (4.82)	5650 (5.28)	4953 (4.63)	5650 (5.28)	4953 (4.63)	5650 (5.28)	4953 (4.63)
Alprazolam N (%)	125251 (1.27)	7298 (6.73)	8801 (8.22)	7006 (6.54)	8801 (8.22)	7006 (6.54)	8801 (8.22)	7006 (6.54)
Levodopa N (%)	44126 (0.45)	3552 (3.28)	4316 (4.03)	3510 (3.28)	4316 (4.03)	3510 (3.28)	4316 (4.03)	3510 (3.28)
Duloxetine N (%)	107184 (1.09)	5692 (5.25)	6463 (6.04)	5442 (5.08)	6463 (6.04)	5442 (5.08)	6463 (6.04)	5442 (5.08)
Zolpidem N (%)	96903 (0.98)	5355 (4.94)	5629 (5.26)	5081 (4.74)	5629 (5.26)	5081 (4.74)	5629 (5.26)	5081 (4.74)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

図 II-3. クエチアピン・糖尿病の2項関係における背景の調整

傾向スコアマッチング法によるNと背景因子の統一前(a)及び統一後(b)。図には調整に用いた疾患・薬物それぞれ200件のうち上位25件のみを示している。

被疑薬・有害事象・併用薬の3項関係の表現方法

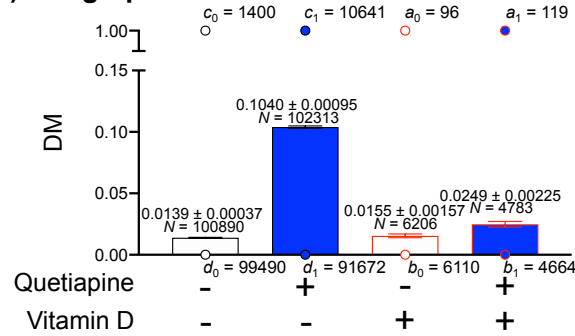
ここまでの検討は被疑薬・有害事象の2項関係の解析手法について、薬剤疫学分野における解析手法と分子薬理学実験における解析手法との対応関係を示しつつ、被疑薬(クエチアピン)の使用と有害事象(糖尿病)報告割合増加との間に強い相関が認められることを確認してきた。しかしながら、クエチアピン使用患者における血糖上昇・糖尿病は既知の有害事象であり、重要なのはそれを再確認することではなく、その解決策を提示することである。図 II-4 以降は被疑薬・有害事象に併用薬を加えた3項関係におけるデータの解析手法に焦点を合わせ、被疑薬使用患者の有害事象を抑制しうる併用薬の探索方法について検討を行う。

図 II-4 は被疑薬(クエチアピン)・有害事象(糖尿病)・併用薬(ビタミン D)の3項関係の表現方法について検討したものである。第一章の表 I-3 では解析対象をクエチアピン使用症例に絞り込んだ上で、各併用薬の有無・糖尿病報告の有無で2×2分割表を作成し併用薬を探索した(図 II-4a の1-2行目に相当)が、分子薬理学実験との対応を考えれば、クエチアピン非使用症例についても同様に解析した4×2分割表を作成し、対照群、クエチアピン使用群、ビタミン D 使用群、クエチアピン・ビタミン D 併用群の4群データを可視化した方が状況を把握しやすくなる(図 II-4b)。図 II-4a の分割表は図 II-1b と同様にダミー変数を用いてカラム型テーブルに変換することで平均値・標準誤差及び N に要約することが可能である(図 II-4c)。

a) Contingency table expression

Table format: Contingency		Outcome A	Outcome B
		DM	No DM
		Y	Y
1	Quetiapine & vitamin D	119	4664
2	Quetiapine & no vitamin D	10641	91672
3	No quetiapine & vitamin D	96	6110
4	No quetiapine & no vitamin D	1400	99490

b) Bar graph visualization



c) Mean, SEM, N summarization

Table format: Grouped		Group A			Group B		
		Quetiapine			No quetiapine		
		Mean	SEM	N	Mean	SEM	N
1	Vitamin D	0.0249	0.00225	4783	0.0155	0.00157	6206
2	No vitamin D	0.1040	0.00095	102313	0.0139	0.00037	100890

図 II-4. クエチアピン・糖尿病・ビタミン D の 3 項関係の表現方法

(a) クエチアピンの傾向スコアマッチング実施後のデータ(図 II-3b)について、併用薬(ビタミン D)の有無で更に分割した 4×2 分割表. (b) 分割表(a)の棒グラフによる可視化. バーは平均値±標準誤差を示している. (c) 分割表(a)の平均値, 標準誤差及び N による要約.

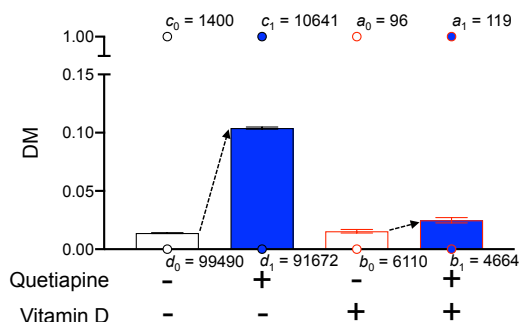
被疑薬・有害事象・併用薬の3項関係における群間の比較

続いて、被疑薬(クエチアピン)及び併用薬(ビタミン D)使用・非使用群における有害事象(糖尿病)報告割合の比較方法について検討する(図 II-5). クエチアピン・ビタミン D 間の相互作用を評価する方法は大きく分けて2種類存在する. 1つは図 II-5a のように、対照群に対するクエチアピン使用群の糖尿病報告割合の増加が、ビタミン D 使用条件下では増加しにくくなることを示す *crossed interaction* であり、もう1つは図 II-5b のように、クエチアピン・ビタミン D 併用群ではビタミン D 非併用群に比べ糖尿病報告割合が減少することを示す *nested interaction* である. 図 II-5c,d の解析結果に示すように、両者は分子薬理学実験において広く用いられている二元配置分散分析とその後の多重比較検定の関係に相当する. 研究[44]では両者を別々に評価しているが、分子薬理学実験では一般に、分散分析により統計学的に有意な差が確認された場合に多重比較法による群間比較を行うのが慣例であり、この対応から言えば *crossed interaction*, *nested interaction* の双方について相互作用を検出できるのが理想的であると言える.

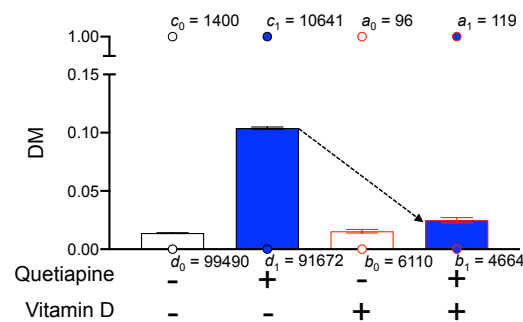
なお、分子薬理学実験では Bonferroni 法や Tukey 法等により検定の多重性を補正した P 値を報告するのが一般的であるが、 P 値が6倍される程度の変化であるため、有害事象ビッグデータを扱う本研究の場合にはほとんど結果に影響しない. 図 II-5c,d の P 値を計算すると図 II-5e,f のようになる.

図 II-2e と同様、一般化線形モデルを用いた解析を適用すれば差のみならず比・オッズ比を指標に計算することや、Wald 法、Score 法、LRT 法による P 値計算が可能となる(図 II-5g,h)ため、併用薬探索に利用可能な指標は多種多様となり、どの指標を用いてどのように探索したのかという推論過程が重要となる. 例えば第一章の表 I-3 における解析ではクエチアピン・併用薬群の $N (a_1+b_1)$ が 1000 以上となる併用薬について網羅的に解析し、*nested interaction* の糖尿病オッズ比が最小となる薬物としてビタミン D を見出しており、研究[8]ではロシグリタゾン・併用薬群の N が 500 以上となる併用薬について網羅的に解析し、*nested interaction* の心筋梗塞オッズ比の Wald 法 P 値が最小となる薬物としてエキセナチドを見出している. 併用薬探索にはこれらの数値データのみならず、臨床的実用性、分子薬理学的妥当性についても考慮した合理的な推論が求められる.

a) Crossed interaction



b) Nested interaction



c) 2-way analysis of variance (ANOVA)

2way ANOVA ANOVA results		
Source of Variation	P value summary	Significant?
Interaction	****	Yes
ANOVA table	F (DFn, DFd)	P value
Interaction	F(1, 214188) = 321.7	P<0.0001
Interaction CI		
Mean diff, A1 - B1	0.009411	
Mean diff, A2 - B2	0.09013	
(A1 - B1) - (A2 - B2)	-0.08072	
95% CI of difference	-0.08954 to -0.07190	

d) Post hoc multiple comparisons

2way ANOVA Multiple comparisons	
Uncorrected Fisher's LSD	
Individual P value	<0.0001
P value summary	****
Significantly different (P < 0.05)?	Yes
One- or two-tailed P value?	Two-tailed
t, df	t=23.47, df=214188
Predicted (LS) mean diff.	-0.07912
95.00% CI of diff.	-0.08573 to -0.07252

e) P value calculation

2-way ANOVA interaction P value = 6.9×10^{-72}

f) P value calculation

Uncorrected Fisher's LSD P value = 1.2×10^{-121}
Bonferroni's corrected P value = 7.4×10^{-121}

g) GLMs for binary data

Crossed interaction		Difference of differences	Ratio of ratios	Ratio of odds ratios
		-0.081	$\times 0.21$	$\times 0.20$
Wald	P	2.7×10^{-167}	1.3×10^{-28}	1.6×10^{-30}
	CI	-0.086 to -0.075	0.16 to 0.28	0.15 to 0.26
Score	P	1.2×10^{-70}	6.6×10^{-34}	1.2×10^{-36}
	CI			
LRT	P	9.0×10^{-93}	2.5×10^{-25}	4.9×10^{-27}
	CI	-0.086 to -0.075	0.16 to 0.28	0.15 to 0.26

h) GLMs for binary data

Nested interaction		Difference	Ratio	Odds ratio
		-0.079	$\times 0.24$	$\times 0.22$
Wald	P	1.5×10^{-229}	1.1×10^{-55}	3.6×10^{-59}
	CI	-0.084 to -0.074	0.20 to 0.29	0.18 to 0.26
Score	P	8.2×10^{-71}	8.2×10^{-71}	8.2×10^{-71}
	CI			
LRT	P	1.7×10^{-96}	1.7×10^{-96}	1.7×10^{-96}
	CI	-0.084 to -0.074	0.20 to 0.28	0.18 to 0.26

図 II-5. クエチアピン・糖尿病・ビタミン D の 3 項関係における群間の比較

(a) Crossed interaction によるビタミン D 併用の影響評価. (b) Nested interaction によるビタミン D 併用の影響評価. (c) 二元配置分散分析の解析結果. (d) 多重比較法の解析結果. (e) 検定(c)の P 値計算結果. (f) 検定(d)の P 値計算結果及び Bonferroni 法による補正結果. (g) 一般化線形モデルによる crossed interaction の効果量, P 値及び 95%信頼区間計算結果. (h) 一般化線形モデルによる nested interaction の効果量, P 値及び 95%信頼区間計算結果. 分子薬理学実験で広く用いられる二元配置分散分析(c,e)とその後の多重比較検定(d,f)に最も類似した傾向の得られる解析法としては差を用いた Score 検定(g,h 黄色部)が該当する. (c)及び(d)は Prism 8, (e-h)は R による解析結果.

被疑薬・有害事象・併用薬の3項関係における背景の調整

続いて、被疑薬(クエチアピン)及び併用薬(ビタミン D)使用・非使用群における背景の調整について検討する(図 II-6)。併用薬解析においても、図 II-3 と同じく傾向スコアマッチング法を適用することで、4群の N と背景因子を技術的に統一することができる。しかしながら、このような二段階のマッチングを経ることで元々1000万件程度あった症例は合計1万件程度(各群 $N = 3,995$)まで大幅に削減されることになる。それ故、本解析手法は FAERS のような大規模有害事象データベースを用いたとしても、あらゆる被疑薬・有害事象の組み合わせに対し有用な併用薬を探索できる訳ではなく、適用可能な研究課題は限られてくることになる。

a) Before matching

b) After matching

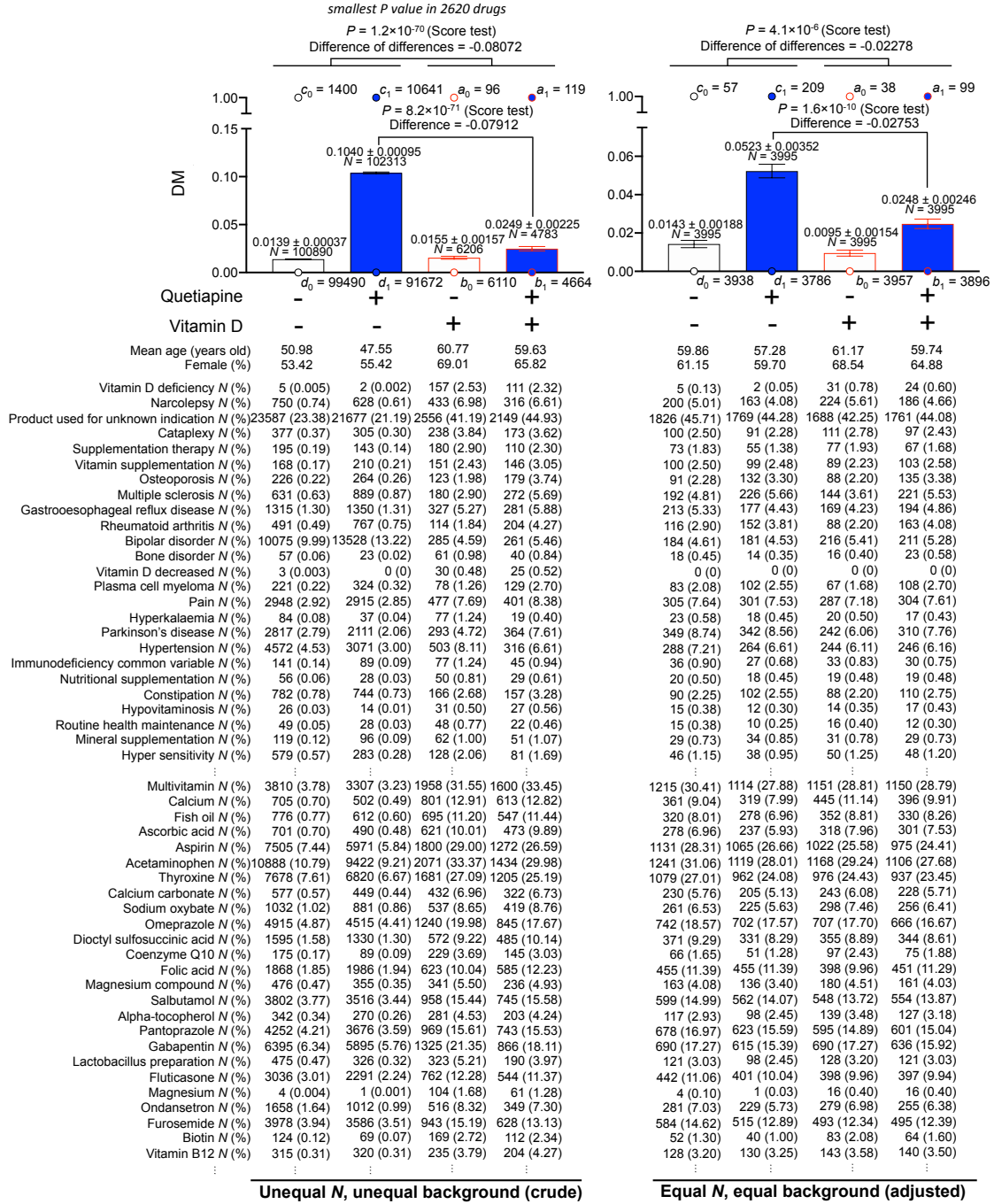


図 II-6. クエチアピン・糖尿病・ビタミン D の 3 項関係における背景の調整
 傾向スコアマッチング法による N と背景因子の統一前(a)及び統一後(b)。図には調整に用いた疾患・薬物それぞれ 200 件のうち上位 25 件(Score 検定 P 値順)のみを示している。

効果量及び検定法に関する薬剤疫学・分子薬理学の慣例の再検討

最後に、図 II-2 で扱った薬剤疫学・分子薬理学の慣例の相違点(5-7)について、

- (i) 有害事象ビッグデータ解析を分子薬理学研究に応用する
- (ii) 被疑薬使用患者の有害事象を抑制する併用薬を探索する

という 2 つの条件から再検討する。

相違点(5)において、有害事象データベースの薬剤疫学的解析では一般にオッズ比が効果量として用いられることを述べた。しかしながら、条件(i)及び(ii)を考慮すると、オッズ比が最適な指標であるとは限らない。オッズ比は数学的対称性から言えば確かに扱いやすい指標であるが、分子薬理学実験では二値データを扱う機会が少なく、平均値の差でも比でも無いオッズ比は分子薬理学研究者にとって理解しづらいという難点がある。加えて、オッズ比を用いた薬物間相互作用解析結果の解釈は単純ではない。例として、被疑薬(クエチアピン)・併用薬(オランザピン)共に有害事象(糖尿病)報告割合増加との相関が認められる場合の薬物間相互作用について検討する(図 II-7a-d)。この場合、オッズ比を指標として相互作用を評価すると、crossed interaction は 0.63 倍に減少する一方、nested interaction は 6.49 倍に増加するため、crossed interaction と nested interaction とで結果が逆転してしまう。一方、差を指標として相互作用を評価した場合、crossed interaction が 0.20 増加、nested interaction が 0.28 増加となるためこのような結果の不一致は生じない。先行研究[45]は比・オッズ比を指標とする積算モデル(multiplicative model)に比べ、差を指標とする加算モデル(additive model)の方が既知の薬物間相互作用を正しく検出する上で有用であると報告している。以上の知見をまとめると、有害事象データベースの薬剤疫学的解析では一般にオッズ比が用いられるが、条件(i)及び(ii)を考慮すると、分子薬理学実験と同じく差を指標とする方が誤解を生じにくいと考えられる。

相違点(6)において、有害事象データベースの薬剤疫学的解析では一般に Wald 法を用いた信頼区間計算が用いられることを述べた。しかしながら、条件(i)及び(ii)を考慮すると、Wald 法も最適な検定法であるとは限らない。Wald 法は計算の容易性から言えば確かに扱いやすい検定法であるが、Wald 法による計算結果は Prism 8 で実行可能な Pearson の χ^2 検定や Fisher の直接確率検定のどちらとも異なるため混乱しやすいという難点がある。加えて、Wald 法を用いた薬物間相互作用解析結果の解釈も単純ではない。例として、被疑薬(クエチアピン)・併用薬(ビタミン D)併用群において有害事象(糖尿病)報告が存在しない($a_1 = 0$)場合の

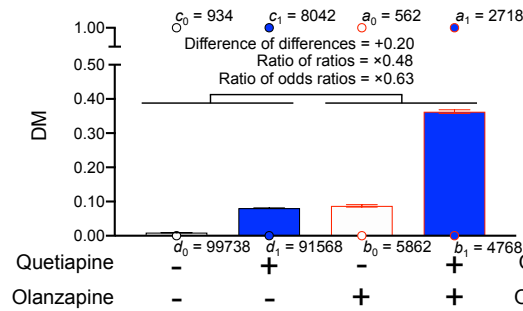
薬物間相互作用について検討する(図 II-7e-h). この場合, Wald 法を用いた解析では P 値が極端に小さい(差)又は大きい(比・オッズ比)結果となり, 統計学的有意性を適切に評価できなくなってしまう. 一方, Score 法を用いた解析ではこのような例においても比較的安定した計算結果を得ることができる. 統計学の専門書[39,40]ではこのような 0 カウントが含まれる場合の解析法として Wald 法よりも Score 法(又は LRT 法)の方が適切であることが述べられている. 以上の知見をまとめると, 有害事象データベースの薬剤疫学的解析では一般に Wald 法が用いられるが, 条件(i)及び(ii)を考慮すると, 分子薬理学実験で広く用いられる二元配置分散分析と多重比較検定に類似した結果の得られる Score 法を用いた方が誤解を生じにくいと考えられる.

相違点(7)において, 疫学分野では一般に P 値よりも効果量とその 95%信頼区間の方が強調されることを述べた. しかしながら, 条件(i)及び(ii)を考慮すると, P 値よりも 95%信頼区間の方がデータの要約法として適切であるとは限らない. 効果量の 95%信頼区間表現ではベースライン値の情報が失われるため, 4 群比較の場合には各群をどのように比較して得られた数値なのかを把握するのが困難となってしまう. また P 値は先行研究[8]のように, 併用薬候補同士を比較し優先順位を付ける上で有用な指標となる. 以上の知見をまとめると, 薬剤疫学分野では一般に効果量とその 95%信頼区間が強調されるが, 条件(i)及び(ii)を考慮すると, 分子薬理学実験と同様に各群の平均値 \pm 標準誤差をグラフにより可視化した上で N 及び P 値を記載し, 必要に応じて効果量の信頼区間情報を付与する方が誤解を生じにくいと考えられる.

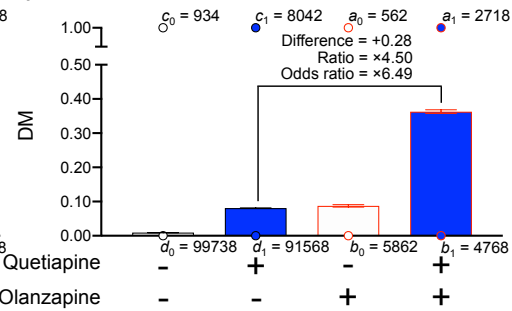
まとめると, 条件(i)及び(ii)を考慮した場合, 薬剤疫学で一般的に用いられているオッズ比・Wald 法・95%信頼区間よりも差・Score 法・ P 値の方が望ましい性質を備えており, 分子薬理学研究者にとっても理解しやすいと考えられる. ただし, これは薬剤疫学分野において従来用いられてきた解析手法を否定するものではなく, 適切な補正・解釈・検証を伴うのであれば, オッズ比や Wald 法, 95%信頼区間を用いた解析でも全く問題は無い. また, Score 法を用いた計算は技術的に高度となってしまうため, 適切な解釈・検証が可能であるならば, 分子薬理学実験で一般に用いられる二元配置分散分析と多重比較検定を 2 値データにそのまま適用しても実用上問題は無いと考えられる.

Example 1. Difference vs. Ratio vs. Odds ratio

a) Crossed interaction



b) Nested interaction



c) GLMs for binary data

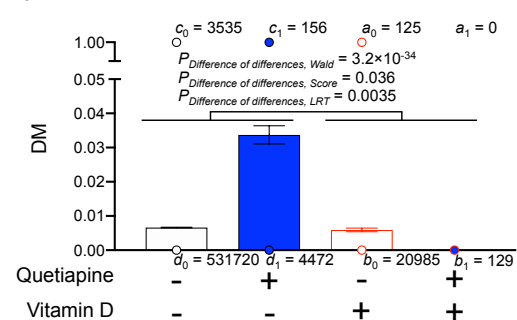
Crossed interaction	Difference of differences +0.20	Ratio of ratios x0.48	Ratio of odds ratios x0.63
Wald	<i>P</i> 3.0×10^{-207} CI 0.19 to 0.22	3.3×10^{-41} 0.43 to 0.53	9.5×10^{-14} 0.56 to 0.71
Score	<i>P</i> 2.7×10^{-184} CI	5.8×10^{-42}	7.4×10^{-14}
LRT	<i>P</i> 6.1×10^{-194} CI 0.19 to 0.22	1.3×10^{-38} 0.43 to 0.53	1.9×10^{-13} 0.56 to 0.72

d) GLMs for binary data

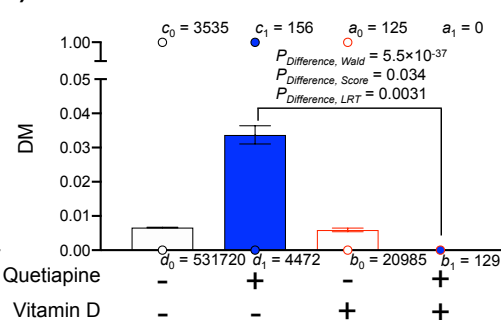
Nested interaction	Difference +0.28	Ratio x4.50	Odds ratio x6.49
Wald	<i>P</i> 10^{-549} CI 0.27 to 0.29	10^{-1410} 4.34 to 4.66	10^{-1067} 6.16 to 6.84
Score	<i>P</i> 10^{-1336} CI	10^{-1336}	10^{-1336}
LRT	<i>P</i> 10^{-903} CI 0.27 to 0.29	10^{-903} 4.33 to 4.66	10^{-903} 6.16 to 6.84

Example 2. Wald vs. Score vs. LRT

e) Crossed interaction



f) Nested interaction



g) GLMs for binary data

Crossed interaction	Difference of differences -0.033	Ratio of ratios x0	Ratio of odds ratios x0
Wald	<i>P</i> 3.2×10^{-34} CI -0.038 to -0.028	1 0 to ∞	1 0 to ∞
Score	<i>P</i> 0.036 CI	0.045	0.045
LRT	<i>P</i> 0.0035 CI	0.0055	0.0055

h) GLMs for binary data

Nested interaction	Difference -0.034	Ratio x0	Odds ratio x0
Wald	<i>P</i> 5.5×10^{-37} CI -0.039 to -0.029	1 0 to ∞	1 0 to ∞
Score	<i>P</i> 0.034 CI	0.034	0.034
LRT	<i>P</i> 0.0031 CI	0.0031	0.0031

図 II-7.2 値データ解析における効果量及び検定法の再検討

例 1. Crossed interaction (a,c)及び nested interaction (b,d)を用いたクエチアピン・オランザピン間の薬物間相互作用評価. 例 2. Crossed interaction (e,g)及び nested interaction (f,h)を用いたクエチアピン・ビタミン D 間の薬物間相互作用評価. 例 2 は JADER データベースを用いた場合の解析結果.

考察

第二章では有害事象データベースの薬剤疫学的解析手法と分子薬理学実験におけるデータ解析との違いを比較しつつ、解析手法の整理・統合を行った。図 II-1,4 における検討では薬剤疫学分野で広く用いられる分割表を分子薬理学実験と同様に平均値、標準誤差及び N に要約できることを示した。図 II-3,6 における検討では傾向スコアマッチング法を用いることで分子薬理学実験と同じく N と背景因子を統一できることを示した。そして図 II-2,5,7 における検討では 2 値データの群間比較法について整理し、分子薬理学研究に応用するには特に差を指標とする Score 検定が有用であることを示した。

薬物有害事象研究をこれまで中心的に担ってきた薬剤疫学分野と動物・細胞実験を中心に発展してきた分子薬理学分野では、扱うデータの性質、表現、解析及び解釈に必要とされる「常識」が大きく異なる。本章では分子薬理学研究者に伝わるよう、できるだけ薬剤疫学研究における専門用語を用いずに両分野における解析手法の対応関係を明示した。本章は有害事象データベースの薬剤疫学的解析手法を、「分子薬理学研究に応用する」という観点から記述されているという点で、新規性・有用性のある内容となっていると考えられる。

FAERS のような臨床有害事象自発報告データベースを分子薬理学研究に応用することは、動物・細胞を用いた分子薬理学実験結果に臨床的有用性を示唆することができるため、「基礎研究において有効であっても臨床で失敗してしまうかもしれない」という懸念を払拭することができる。加えて、FAERS は実臨床において報告する価値があると判断された薬物有害事象を蓄積したデータベースであることから、報告数の多さは潜在的な臨床ニーズの高さを直接的に示しており、研究の意義をデータで直接示すことが可能である。更に、FAERS は他の多くの薬剤疫学研究データベースと異なり Web 上に一般公開されているため、解析結果の再現性や推論の合理性を他研究者が追跡確認できるという利点がある。これらの臨床的有用性、ニーズ志向性、研究の透明性という観点から、有害事象ビッグデータ解析を分子薬理学研究に応用することは基礎研究と臨床研究の間の溝を埋める有効な戦略となると考えられる。

一方で、臨床有害事象自発報告データベースを用いた解析には同時に多くの限界も存在する。第一に、自発報告データベースでは薬物使用患者全体(分母)の情報や詳細な時系列情報を入手することができないため、有害事象のリスクや実態を正確に求めるには不十分であるという点が挙げられる。そのため、FAERS

のような自発報告データベースにおいて検出された仮説は、レセプトや電子カルテといった他の薬剤疫学研究データベースを用いて強化することが望ましい[46]。このように、独立した臨床データベースを複数組み合わせることにより、臨床的意義をより適切に示すことができるようになると考えられる。第二に、臨床データベースの後方視的解析は相関関係を示すことはできても因果関係を証明するには不十分であるという点が挙げられる。傾向スコアマッチングのような優れた解析手法を適用したとしても、予想外の交絡の影響まで完全に排除できるとは限らない。そのため、臨床データベースの後方視的解析により得られた仮説は、動物・細胞を用いた前方視的実験により検証することが望ましい[46]。このように、臨床データベース解析と動物・細胞実験による検証を組み合わせることにより、薬物有害事象の因果関係をより適切に示すことができるようになると考えられる。

有害事象ビッグデータを用いた研究の先駆者である N. P. Tatonetti は次世代型の薬物安全性研究手法として、検出(detection)、強化(corroboration)、検証(validation)という上述の3段階を踏むことが新たな薬効及び薬物間相互作用を見出す上で重要であると主張している[46]。有害事象自発報告データベースの解析により検出された薬物の有用性について、電子カルテ・レセプトデータベース解析と組み合わせることにより強化した研究[8,47-49]や、動物・細胞実験と組み合わせることにより検証した研究[8,50-52]は近年少しずつ増加しているが、複数領域に跨がる学際的な研究となるためその手法はまだ確立されておらず、各研究者がそれぞれ試行錯誤しながら進めている状況にある。本章の内容はそのような状況を整理し、有害事象データベースの薬剤疫学的解析を分子薬理学研究の世界にスムーズに導入できるよう、両分野における解析手法の対応関係をまとめたものである。本章の研究成果により、新たな分子薬理学研究ツールとしての有害事象ビッグデータ解析の活用促進が期待される。

総括及び結論

本研究において、著者は分子薬理学研究分野における有害事象ビッグデータの活用可能性について検討を行い、以下の新知見を得た。

第一章では、有害事象ビッグデータ解析の活用例として非定型抗精神病薬(クエチアピン)使用に伴う高血糖・糖尿病に注目し、FAERS データベースの解析によりビタミン D 製剤の併用という新たな解決策を見出した。また、この仮説に対し動物・細胞を用いた薬理学的実験による検証を行い、クエチアピン使用に伴う血糖上昇・インスリン抵抗性がビタミン D の併用により確かに抑制されることを実証した。加えて、クエチアピン・糖尿病・ビタミン D という 3 項の共通点から、糖尿病症状における *Pik3r1* 遺伝子を介した新たな分子メカニズムの関与を明らかにした。

第二章では、有害事象ビッグデータの薬剤疫学的解析を分子薬理学研究ツールとして確立することを目指し解析手法の整理・統合を試みた。データの表現方法として、被疑薬・有害事象・併用薬の 3 項関係を分子薬理学実験と同様に平均値、標準誤差及び N に要約し可視化できることを示した。また、背景の調整方法として、傾向スコアマッチング法の適用により分子薬理学実験と同様に各群の N と背景因子を統一した結果が得られることを示した。加えて、群間の比較方法として、差を指標とする Score 検定法が分子薬理学実験において一般に用いられる二元配置分散分析及び多重比較法に相当する有用な手法であることを示した。

以上、著者は本研究において、有害事象ビッグデータを用いた薬剤疫学的解析が、薬物有害事象の新たな解決策の提示、既存薬の新たな薬効の探索、病態の新たな分子メカニズムの解明に有用であることを示し、本手法を新たな分子薬理学研究ツールとして活用できるよう解析手法の整理・統合を行った。本研究成果を踏まえ、分子薬理学研究に臨床副作用自発報告ビッグデータを活用していくことで、分子から臨床までを統合する新たな研究スタイルの確立が期待される。

謝辞

本研究の遂行に際し、終始暖かい支援とご指導、ご鞭撻を賜りました京都大学大学院薬学研究科金子周司教授に謹んで感謝の意を表します。また、直接ご指導を賜り、数多くの有益なご助言を頂きました京都大学医学部附属病院薬剤部中川貴之准教授、京都大学大学院薬学研究科白川久志准教授並びに永安一樹助教に心から感謝の意を表します。更に、終始有益なご助言を頂きました京都大学赤池昭紀名誉教授、富山大学大学院医学薬学研究部久米利明教授、神戸薬科大学泉安彦講師に深く感謝申し上げます。研究生活を始めるにあたり、ご指導、ご助言を頂きました崎元伸哉博士、崎元萌博士、宗可奈子博士、土岸広治博士、勇昂一博士、西谷直也博士、三宅崇仁博士、浅岡希美博士、宮之原遵博士をはじめ、諸先輩方に心から感謝致します。そして、同じ研究チームにおいて互いに切磋琢磨し合えた古田晴香学士、長岡巧樹修士、戸田千尋さん、Soni Siswanto 修士、樋口拓哉修士、周子剣学士、岩井信太郎君、山本浩貴学士、栢沼玄君、小林萌音さんに心からお礼を述べると共に、京都大学大学院生体機能解析学分野及び京都大学大学院薬品作用解析学分野の諸氏に深く感謝致します。最後に、著者の良き理解者であり研究活動に専念できるよう終始支えてくださった父直之、母智春、兄寛人、兄利行、弟大和、及び祖母松永栄子に心から感謝致します。

発表論文目録

第一章

非定型抗精神病薬による高血糖に対する解決策の提示及び分子機序の探索

Nagashima Takuya, Shirakawa Hisashi, Nakagawa Takayuki, Kaneko Shuji.

Scientific reports (2016) 6:26375 に掲載

第二章

分子薬理学研究ツールとしての有害事象ビッグデータの解析手法開発

未発表

参考文献

1. American Diabetes Association, American Psychiatric Association, American Association of Clinical Endocrinologists & North American Association for the Study of Obesity. Consensus development conference on antipsychotic drugs and obesity and diabetes. *Diabetes Care* **2**, 596–601 (2004).
2. Sernyak, M. J. *et al.* Association of diabetes mellitus with use of atypical neuroleptics in the treatment of schizophrenia. *Am. J. Psychiatry* **159**, 561–566 (2002).
3. Ely, S. F., Neitzel, A. R. & Gill, J. R. Fatal diabetic ketoacidosis and antipsychotic medication. *J. Forensic Sci.* **58**, 398–403 (2013).
4. Buchholz, S., Morrow, A. F. & Coleman, P. L. Atypical antipsychotic-induced diabetes mellitus: an update on epidemiology and postulated mechanisms. *Intern. Med. J.* **38**, 602–606 (2008).
5. Rojo, L. E. *et al.* Metabolic syndrome and obesity among users of second generation antipsychotics: A global challenge for modern psychopharmacology. *Pharmacol. Res.* **101**, 74–85 (2015).
6. Tatonetti, N. P. *et al.* Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels. *Clin. Pharmacol. Ther.* **90**, 133–142 (2011).
7. Tatonetti, N. P., Ye, P. P., Daneshjou, R. & Altman, R. B. Data-driven prediction of drug effects and interactions. *Sci. Transl. Med.* **4**, 125ra31 (2012).
8. Zhao, S. *et al.* Systems pharmacology of adverse event mitigation by drug combinations. *Sci. Transl. Med.* **5**, 206ra140 (2013).
9. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
10. Kanehisa, M. *et al.* KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44** (Database issue), D457–462 (2016).
11. Ganter, B., Snyder, R. D., Halbert, D. N. & Lee, M. D. Toxicogenomics in drug discovery and development: mechanistic analysis of compound/class-dependent effects using the DrugMatrix database. *Pharmacogenomics* **7**, 1025–1044 (2006).

12. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002).
13. Barrett, T. *et al.* NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Res.* **41** (Database issue), D991–995 (2013).
14. Reaven, G. M. Banting lecture 1988. Role of insulin resistance in human disease. *Diabetes* **37**, 1595–1607 (1988).
15. Barroso, I. *et al.* Candidate gene association study in type 2 diabetes indicates a role for genes involved in beta-cell function as well as insulin action. *PLoS Biol.* **1**, E20 (2003).
16. Thauvin-Robinet, C. *et al.* *PIK3R1* mutations cause syndromic insulin resistance with lipodystrophy. *Am. J. Hum. Genet.* **93**, 141–149 (2013).
17. DeFronzo, R. A. & Tripathy, D. Skeletal muscle insulin resistance is the primary defect in type 2 diabetes. *Diabetes Care* **32**, S157–193 (2009).
18. Maestro, B., Dávila, N., Carranza, M. C. & Calle, C. Identification of a Vitamin D response element in the human insulin receptor gene promoter. *J. Steroid Biochem. Mol. Biol.* **84**, 223–230 (2003).
19. Nwosu, B. U. *et al.* A potential role for adjunctive vitamin D therapy in the management of weight gain and metabolic side effects of second-generation antipsychotics. *J. Pediatr. Endocrinol. Metab.* **24**, 619–626 (2011).
20. Thakurathi, N. *et al.* Open-label pilot study on vitamin D3 supplementation for antipsychotic-associated metabolic anomalies. *Int. Clin. Psychopharmacol.* **28**, 275–282 (2013).
21. Dang, R. *et al.* Vitamin D deficiency exacerbates atypical antipsychotic-induced metabolic side effects in rats: involvement of the INSIG/SREBP pathway. *Eur. Neuropsychopharmacol.* **25**, 1239–1247 (2015).
22. Pittas, A. G., Lau, J., Hu, F. B. & Dawson-Hughes, B. The role of vitamin D and calcium in type 2 diabetes. A systematic review and meta-analysis. *J. Clin. Endocrinol. Metab.* **92**, 2017–2029 (2007).
23. Miller, D. D. Review and management of clozapine side effects. *J. Clin. Psychiatry* **61**, 14–17 (2000).
24. Haussler, M. R. *et al.* Molecular mechanisms of vitamin D action. *Calcif. Tissue Int.* **92**, 77–98 (2013).

25. Luo, J. *et al.* Loss of class IA PI3K signaling in muscle leads to impaired muscle growth, insulin response, and hyperlipidemia. *Cell Metab.* **3**, 355–366 (2006).
26. Thong, F. S., Dugani, C. B. & Klip, A. Turning signals on and off: GLUT4 traffic in the insulin-signaling highway. *Physiology (Bethesda)* **20**, 271–284 (2005).
27. Engl, J. *et al.* Olanzapine impairs glycogen synthesis and insulin signaling in L6 skeletal muscle cells. *Mol. Psychiatry* **10**, 1089–96 (2005).
28. Panariello, F. *et al.* Clozapine impairs insulin action by up-regulating Akt phosphorylation and Ped/Pea-15 protein abundance. *J. Cell. Physiol.* **227**, 1485–1492 (2012).
29. Zhou, Q. G. *et al.* 1,25-Dihydroxyvitamin D improved the free fatty-acid-induced insulin resistance in cultured C2C12 cells. *Diabetes Metab. Res. Rev.* **24**, 459–464 (2008).
30. Ngai, Y. F. *et al.* Quetiapine treatment in youth is associated with decreased insulin secretion. *J. Clin. Psychopharmacol.* **34**, 359–364 (2014).
31. Ozasa, R. *et al.* The antipsychotic olanzapine induces apoptosis in insulin-secreting pancreatic β cells by blocking PERK-mediated translational attenuation. *Cell Struct. Funct.* **38**, 183–195 (2013).
32. Lorberbaum, T. *et al.* Systems pharmacology augments drug safety surveillance. *Clin. Pharmacol. Ther.* **97**, 151–158 (2015).
33. Kim, H. *et al.* Meta-analysis of large-scale toxicogenomic data finds neuronal regeneration related protein and cathepsin D to be novel biomarkers of drug-induced toxicity. *PLoS One* **10**, e0136698 (2015).
34. Sakaeda, T., Tamon, A., Kadoyama, K. & Okuno, Y. Data mining of the public version of the FDA adverse event reporting system. *Int. J. Med. Sci.* **10**, 796–803 (2013).
35. Kondo, M. A. *et al.* Unique pharmacological actions of atypical neuroleptic quetiapine: possible role in cell cycle/fate control. *Transl. Psychiatry* **3**, e243 (2013).
36. Enea, M. *et al.* speedglm: Fitting linear and generalized linear models to large data sets. R package version 0.3-2 (2017).
37. Sekhon, J. S. Multivariate and propensity score matching software with automated balance optimization: the Matching package for R. *Journal of Statistical Software* **42**, 1–52 (2011).

38. Ho, D. E. *et al.* MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software* **42**, 1–28 (2011).
39. Agresti, A. An introduction to categorical data analysis, 3rd edition. John Wiley & Sons (2018).
40. Agresti, A. Categorical data analysis, 3rd edition. John Wiley & Sons (2013).
41. Rothman, K., Greenland, S. & Lash, TL. Modern epidemiology, 3rd edition. Lippincott Williams & Wilkins (2008).
42. Perry, B. I. *et al.* The association between first-episode psychosis and abnormal glycaemic control: systematic review and meta-analysis. *Lancet Psychiatry* **3**, 1049–1058 (2016).
43. Xu, D. *et al.* MSBIS: A multi-step biomedical informatics screening approach for identifying medications that mitigate the risks of metoclopramide-induced tardive dyskinesia. *EBioMedicine* **26**, 132–137 (2017).
44. Li, Y., Zhang P., Sun Z. & Hu J. Data-driven prediction of beneficial drug combinations in spontaneous reporting systems. *AMIA Annu. Symp. Proc.*, 808–817 (2016).
45. Thakrar, B. T., Grundschober, S. B., & Doessegger, L. Detecting signals of drug-drug interactions in a spontaneous reports database. *Br. J. Clin. Pharmacol.* **64**, 489–495 (2007).
46. Tatonetti, N. P. The next generation of drug safety science: coupling detection, corroboration, and validation to discover novel drug effects and drug-drug interactions. *Clin. Pharmacol. Ther.* **103**, 177–179 (2018).
47. Takada, M., Fujimoto, M., Motomura, H. & Hosomi, K. Inverse association between sodium channel-blocking antiepileptic drug use and cancer: data mining of spontaneous reporting and claims databases. *Int. J. Med. Sci.* **13**, 48–59 (2016).
48. Hosomi, K. *et al.* An integrative approach using real-world data to identify alternative therapeutic uses of existing drugs. *PLoS One* **13**, e0204648 (2018).
49. Yokoyama, S. *et al.* Integrative analysis of clinical and bioinformatics databases to identify anticancer properties of digoxin. *Sci. Rep.* **9**, 16597 (2019).
50. Hashikawa, N. *et al.* HSP105 prevents depression-like behavior by increasing hippocampal brain-derived neurotrophic factor levels in mice. *Sci. Adv.* **3**, e1603014 (2017).
51. Horinouchi, Y. *et al.* Renoprotective effects of a factor Xa inhibitor: fusion of

- basic research and a database analysis. *Sci. Rep.* **8**, 10858 (2018).
52. Izawa-Ishizawa, Y. *et al.* Development of a novel aortic dissection mouse model and evaluation of drug efficacy using in-vivo assays and database analyses. *J. Hypertens.* **37**, 73–83 (2019).