



(12) 发明专利申请

(10) 申请公布号 CN 114250288 A

(43) 申请公布日 2022.03.29

(21) 申请号 202111281022.6

G16H 50/30 (2018.01)

(22) 申请日 2021.11.01

G16H 50/50 (2018.01)

(71) 申请人 苏州市广济医院

G06K 9/62 (2022.01)

地址 215131 江苏省苏州市广前路11号

G06N 20/00 (2019.01)

(72) 发明人 张晓斌 杜向东 田晴 殷光中
朱峰 王传跃

(74) 专利代理机构 北京预立生科知识产权代理
有限公司 11736

代理人 李红伟 孟祥斌

(51) Int. Cl.

G12Q 1/6883 (2018.01)

G01H 17/00 (2006.01)

G16B 5/20 (2019.01)

G16B 15/00 (2019.01)

G16H 50/20 (2018.01)

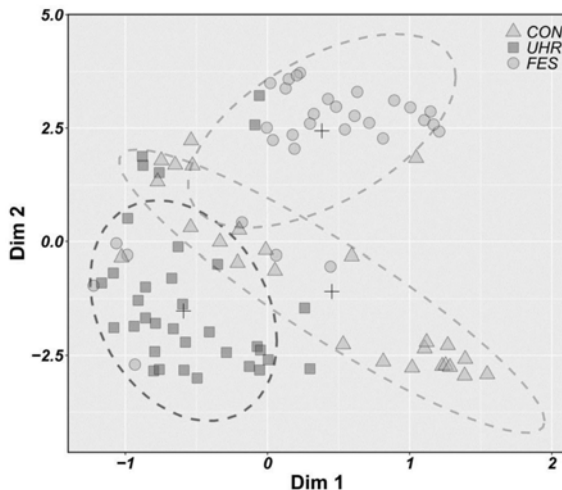
权利要求书6页 说明书16页 附图6页

(54) 发明名称

DNA甲基化特征和前脉冲抑制特征在精神分裂症诊断中的应用

(57) 摘要

本发明公开了DNA甲基化特征和前脉冲抑制特征在精神分裂症诊断中的应用,本发明首次将DNA甲基化特征和前脉冲抑制特征用作临床生物标志物,联合应用于诊断鉴别首发精神分裂症、超高危人群和健康人群,具有客观、准确、有效等优点,在临床上具有非常好的应用前景,为精神分裂症的早期识别和干预奠定了基础。



1. DNA甲基化特征、和/或前脉冲抑制特征作为生物标志物在制备诊断精神分裂症的产品中的应用。

2. 根据权利要求1所述的应用,其特征在于,所述应用包括DNA甲基化特征、和/或前脉冲抑制特征作为生物标志物在制备诊断首发精神分裂症的产品中的应用;

优选地,所述DNA甲基化特征包括900个DNA甲基化特征位点;

更优选地,所述900个DNA甲基化特征位点包括cg25224760、cg25122934、cg25783352、cg00693060、cg23195763、cg25176297、cg18516940、cg22030766、cg21163478、cg15153114、cg03515346、cg16210324、cg18561199、cg24344143、cg03803284、cg23243856、cg27400655、cg13748287、cg14044124、cg12390678、cg18889956、cg07746811、cg25799064、cg15796941、cg24727290、cg10924691、cg11631427、cg26369259、cg22594586、cg01511844、cg15782771、cg18065318、cg25672411、cg23147317、cg18718234、cg15430422、cg22910549、cg06759901、cg25242729、cg22879458、cg25920734、cg25763426、cg20592447、cg04570624、cg20793712、cg25929589、cg04008557、cg10294363、cg23879228、cg12904004、cg04584468、cg24025893、cg17173330、cg01758041、cg17681524、cg24399801、cg07745707、cg25567990、cg10164186、cg21686797、cg17250082、cg13054419、cg25122125、cg13384409、cg27475680、cg26619156、cg24982343、cg17391741、cg11080430、cg24981400、cg19940644、cg20781380、cg20901993、cg13850825、cg02160684、cg24365551、cg22027671、cg16541340、cg15126273、cg21239409、cg18003231、cg24766555、cg25178683、cg21351992、cg11508828、cg24913623、cg16522885、cg18411015、cg22930950、cg17058296、cg27175294、cg19129200、cg12285326、cg11321371、cg24172989、cg20927425、cg10482356、cg04739228、cg14892768、cg22790758、cg02860199、cg11688731、cg20795889、cg06107159、cg17877898、cg16781264、cg06508379、cg15525341、cg27389290、cg27613174、cg26095809、cg10733080、cg14427074、cg13245883、cg24371156、cg16356539、cg01997461、cg17036570、cg23464565、cg27259382、cg24554865、cg26473110、cg10120112、cg23856267、cg05283486、cg25818522、cg24451839、cg21362233、cg09311571、cg19136704、cg20100445、cg17365824、cg22007638、cg14478036、cg27124109、cg01367037、cg23976937、cg00419512、cg13918430、cg14700479、cg14617167、cg26765871、cg13088598、cg02172133、cg19011911、cg17195231、cg26315559、cg11623950、cg26202181、cg21722655、cg06065611、cg26124670、cg27652200、cg01573019、cg17830740、cg19506491、cg26030037、cg17717125、cg22643214、cg17569754、cg11202345、cg20471798、cg26686249、cg18944640、cg22064182、cg17215601、cg27289295、cg02237470、cg26862091、cg13823585、cg09315290、cg23180925、cg24524946、cg19733221、cg24748771、cg22894329、cg13042714、cg23563927、cg08134068、cg25418474、cg20965876、cg25139963、cg20851097、cg10554947、cg27022602、cg25603927、cg12571055、cg10555383、cg04950413、cg07768486、cg26110834、cg21059366、cg24132991、cg19854991、cg26616258、cg23671319、cg04538470、cg07343706、cg23892310、cg22871908、cg02677877、cg24694879、cg01699425、cg24953213、cg27465618、cg25267487、cg10851774、cg02348488、cg18885125、cg25739742、cg00064261、cg09694403、cg18063493、cg15049549、cg23881601、cg08331138、cg23201527、cg26916621、cg22933195、cg22908033、cg04190450、cg24693635、cg17138984、cg16611213、cg09500174、cg09439920、cg15028904、cg24247231、cg06520003、cg00356723、cg27643917、cg25902889、cg15709265、cg22079361、

cg14609682、cg21518713、cg11002709、cg17031944、cg01811230、cg21378920、cg19362632、
cg16260889、cg09319649、cg08057037、cg02100464、cg27272679、cg14880340、cg24013954、
cg24957704、cg22013228、cg18301955、cg10452175、cg05587552、cg23572751、cg05156242、
cg06884029、cg07108214、cg23847843、cg12858166、cg11868809、cg15282731、cg05393023、
cg02384255、cg03538137、cg12265130、cg13848607、cg16254946、cg27141176、cg06045225、
cg11568925、cg12062504、cg10926330、cg26930230、cg22332577、cg24836607、cg02098075、
cg26924679、cg14650228、cg11384205、cg00847359、cg13525067、cg16482175、cg13036702、
cg19120496、cg13857168、cg22413912、cg09358961、cg05328885、cg23062115、cg20563131、
cg18394235、cg09756436、cg06642177、cg19308620、cg20728314、cg01142417、cg19802477、
cg01072259、cg08722290、cg22910769、cg10117603、cg00551487、cg08314949、cg09941712、
cg06525127、cg01035812、cg17942689、cg19172429、cg05842490、cg18594106、cg05612346、
cg17389077、cg24964635、cg25621420、cg02203528、cg17654766、cg24957609、cg20296990、
cg17426070、cg14776168、cg04744810、cg22270828、cg00934864、cg04756223、cg18116174、
cg10286969、cg04641380、cg21778244、cg19251740、cg16619193、cg26837906、cg19229071、
cg19851029、cg01858657、cg06202930、cg10588705、cg04900856、cg10935762、cg26220033、
cg26259271、cg15560837、cg18793005、cg15663074、cg27159443、cg07134804、cg19977871、
cg26945392、cg08040115、cg04855546、cg18038737、cg17714225、cg08638514、cg27173374、
cg26643617、cg16993936、cg01894875、cg07349169、cg19988408、cg08205921、cg00459299、
cg05061471、cg14648186、cg11812015、cg14504291、cg21040513、cg16629408、cg09370941、
cg18328477、cg18266383、cg17876742、cg17730591、cg11775190、cg17343033、cg07340025、
cg07072638、cg10390228、cg14998917、cg19564077、cg12251779、cg03024403、cg13719628、
cg15137408、cg17335199、cg09650487、cg12298697、cg12468966、cg09603833、cg04369787、
cg16467919、cg21544091、cg11488893、cg21896553、cg02812534、cg18443412、cg08501989、
cg24257454、cg10880459、cg05665447、cg12331332、cg16937902、cg16316042、cg04535320、
cg13996738、cg00608943、cg13676583、cg18160999、cg21675569、cg01217923、cg13651234、
cg25197240、cg21548152、cg01772073、cg14382976、cg24241688、cg13426198、cg01841782、
cg17936572、cg16178855、cg00369126、cg15381380、cg13174660、cg15851317、cg12401644、
cg24365633、cg03828570、cg11590902、cg24052284、cg13692650、cg05064489、cg24055461、
cg13789353、cg00234363、cg16385933、cg09888026、cg07957070、cg18582505、cg19132701、
cg12808194、cg10121816、cg19691425、cg01926051、cg19324027、cg16470919、cg12320972、
cg02131013、cg08774990、cg06050385、cg10967622、cg15937259、cg18623931、cg19484548、
cg11002033、cg05682017、cg14429240、cg05197515、cg01438434、cg08877624、cg22252041、
cg05637384、cg17953869、cg10718940、cg22138494、cg13565624、cg06647133、cg10220104、
cg26291385、cg10249373、cg22044892、cg12483005、cg27229471、cg25370362、cg02477579、
cg06468646、cg02414586、cg02483627、cg13505944、cg09084892、cg07237023、cg00232772、
cg09200468、cg02581147、cg01428397、cg10157558、cg08464954、cg21403762、cg06037240、
cg05919900、cg00867317、cg02215945、cg14473016、cg11297423、cg04305319、cg07322293、
cg18793773、cg15600512、cg07870479、cg16278747、cg11605394、cg13096007、cg08091867、
cg24913501、cg03224013、cg00169917、cg20695587、cg02606535、cg16353907、cg02968116、

cg05976544、cg05236879、cg09721840、cg07599136、cg15887150、cg03005121、cg08510930、
cg01072942、cg19212949、cg09528372、cg01831459、cg15044205、cg08967338、cg07848764、
cg03644971、cg02894527、cg01360325、cg00700866、cg18504368、cg05141656、cg02263704、
cg00524443、cg23402560、cg09340485、cg18279180、cg17424473、cg15701419、cg17509462、
cg08845028、cg12290190、cg17084361、cg15405625、cg00723961、cg16183122、cg06538872、
cg03217733、cg22168617、cg07044458、cg21159260、cg08096851、cg03705571、cg09099893、
cg20226150、cg21455883、cg07938869、cg20336788、cg06814654、cg20643659、cg05271612、
cg00576340、cg12161959、cg11369706、cg04425202、cg11703939、cg03114004、cg01705612、
cg11706083、cg12388760、cg01771651、cg10701640、cg10691323、cg09414156、cg01627046、
cg17344560、cg03986909、cg09936757、cg01606576、cg05933932、cg15646610、cg07066790、
cg06458679、cg15565231、cg04734610、cg04526750、cg04290747、cg03204444、cg14453769、
cg14500267、cg13065504、cg03765047、cg12888632、cg00389386、cg14843586、cg12402427、
cg10703338、cg12394910、cg11036189、cg13515074、cg10440450、cg20355084、cg10075524、
cg08071164、cg07678448、cg08193100、cg05016746、cg05580671、cg05718343、cg07431064、
cg02843332、cg08741842、cg05497345、cg02540504、cg01582438、cg00369658、cg04181189、
cg02404489、cg00338113、cg00770663、cg00031362、cg00082739、cg00107488、cg00203089、
cg00223046、cg00230381、cg00321703、cg00417419、cg00459043、cg00480856、cg00540722、
cg00578363、cg00713465、cg00771653、cg00787055、cg00814990、cg00838379、cg00934564、
cg00948102、cg01102220、cg01148568、cg01158804、cg01261775、cg01372572、cg01385063、
cg01422467、cg01479738、cg01508386、cg01566396、cg01585372、cg01727923、cg02039058、
cg02068505、cg02173970、cg02278803、cg02292206、cg02447937、cg02481842、cg02484210、
cg02545393、cg02641277、cg02642822、cg02650286、cg02685016、cg02708898、cg02724747、
cg02741655、cg02794096、cg02800362、cg02803629、cg02873991、cg02916932、cg02921122、
cg03031823、cg03110787、cg03123782、cg03131097、cg03145322、cg03201507、cg03223467、
cg03303025、cg03447137、cg03505995、cg03550508、cg03553910、cg03600697、cg03608224、
cg03619332、cg03643709、cg03741350、cg03764027、cg03799192、cg03847642、cg03849780、
cg03864121、cg03880355、cg03892062、cg04001842、cg04059695、cg04101729、cg04113225、
cg04201335、cg04201365、cg04251616、cg04294058、cg04477962、cg04514292、cg04555107、
cg04619882、cg04791145、cg04832450、cg04833648、cg04843801、cg04889960、cg04925085、
cg04939919、cg04962756、cg04967200、cg05031519、cg05116088、cg05119778、cg05221370、
cg05269534、cg05282641、cg05333753、cg05341567、cg05369351、cg05399718、cg05470179、
cg05529874、cg05571437、cg05572751、cg05607935、cg05625951、cg05659199、cg05899224、
cg05949034、cg05971212、cg05977964、cg05978571、cg06100147、cg06185734、cg06250386、
cg06307176、cg06508886、cg06527318、cg06550214、cg06588466、cg06589239、cg06613034、
cg06627532、cg06760899、cg06807379、cg06813554、cg06909254、cg07008701、cg07078225、
cg07119973、cg07176514、cg07177889、cg07245558、cg07310500、cg07346747、cg07372520、
cg07403258、cg07618581、cg07747241、cg07754486、cg07850987、cg07936305、cg07939245、
cg07972159、cg08088948、cg08104023、cg08141194、cg08305707、cg08362804、cg08371190、
cg08578136、cg08611508、cg08616234、cg08791782、cg08817937、cg08864042、cg08894540、

cg08915378、cg08944086、cg08951822、cg08955275、cg09050775、cg09147024、cg09163921、cg09173939、cg09250965、cg09262552、cg09382942、cg09444206、cg09460462、cg09465855、cg09514717、cg09524686、cg09528462、cg09615620、cg09619883、cg09631044、cg09678035、cg09682198、cg09694764、cg09726046、cg09747638、cg09802873、cg09806262、cg09983708、cg10033694、cg10052615、cg10158249、cg10163338、cg10169976、cg10214132、cg10247383、cg10343071、cg10488141、cg10488637、cg10549357、cg10587782、cg10589249、cg10730208、cg10758986、cg10803722、cg10841956、cg10844118、cg10973146、cg10989690、cg11021611、cg11029191、cg11046887、cg11051139、cg11187110、cg11251399、cg11466109、cg11466321、cg11585769、cg11611011、cg11625897、cg11673867、cg11709065、cg11770888、cg11836369、cg11854579、cg11873026、cg11995137、cg11999199、cg12005153、cg12092351、cg12099459、cg12105770、cg12107522、cg12154434、cg12192122、cg12228478、cg12332083、cg12470114、cg12472085、cg12500602、cg12507168、cg12508336、cg12553110、cg12595444、cg12666921、cg12832751、cg12841061、cg12884605、cg13013644、cg13248789、cg13254556、cg13282635、cg13318914、cg13413789、cg13435263、cg13443092、cg13455586、cg13456654、cg13471352、cg13476139、cg13621410、cg13629807、cg13647660、cg13708908、cg13714987、cg13728439、cg13731422、cg13749070、cg13779963、cg13807254、cg13962664、cg13962718、cg13994177、cg13999554、cg14019951、cg14036868、cg14122836、cg14125353、cg14131316、cg14221825、cg14280583;

优选地,所述前脉冲抑制特征包括肌电变化百分比;

更优选地,所述肌电变化百分比(%)的计算公式为:肌电变化百分比(%) = (1-S2/S1) × 100%;

其中,S1为仅含惊刺激的前脉冲抑制的肌电幅值,S2为同时含前刺激的前脉冲抑制的肌电幅值;

最优选地,所述惊刺激为白噪声,持续时间40ms,声压级100dB;

最优选地,所述前刺激为白噪声,分为左声道或右声道领先3ms,持续时间150ms,声压级65dB。

3. 根据权利要求1所述的应用,其特征在于,所述应用包括DNA甲基化特征、和/或前脉冲抑制特征作为生物标志物在制备诊断精神分裂症超高危人群的产品中的应用;

优选地,所述DNA甲基化特征包括4个DNA甲基化特征位点;

更优选地,所述4个DNA甲基化特征位点包括cg25376875、cg27415006、cg09543727、cg00213281;

优选地,所述前脉冲抑制特征为权利要求2中所述的前脉冲抑制特征。

4. 根据权利要求1所述的应用,其特征在于,所述应用包括DNA甲基化特征、和/或前脉冲抑制特征作为生物标志物在制备鉴别诊断首发精神分裂症患者、精神分裂症超高危人群、健康人群的产品中的应用;

优选地,所述DNA甲基化特征包括10个DNA甲基化特征位点;

更优选地,所述10个DNA甲基化特征位点包括cg01511844、cg23076086、cg01807407、cg15782771、cg04538470、cg04218099、cg10297617、cg14240820、cg25929589、cg10315533;

优选地,所述前脉冲抑制特征为权利要求2中所述的前脉冲抑制特征。

5. 一种用于评估待测个体患首发精神分裂症风险的诊断装置,其特征在於,所述诊断装置包括:

(1) 检测单元,该检测单元包括检测待测个体DNA甲基化特征的试剂、和/或检测待测个体前脉冲抑制特征的设备 and 声音素材;

优选地,所述待测个体为未用药者;

优选地,所述DNA甲基化特征为权利要求2中所述的DNA甲基化特征;

优选地,所述前脉冲抑制特征为权利要求2中所述的前脉冲抑制特征;

(2) 分析单元,该分析单元用于对检测单元所得的检测结果进行分析,评估待测个体首发精神分裂症的患病风险;

优选地,所述分析单元还包括采用随机森林算法建模得到包含DNA甲基化特征的DNA甲基化模型、和/或包含前脉冲抑制特征的前脉冲抑制模型;

更优选地,所述分析单元还包括对DNA甲基化模型、前脉冲抑制模型采用10折交叉验证得出最终模拟值,并基于所述最终模拟值构建包含DNA甲基化特征和前脉冲抑制特征的组合模型。

6. 一种用于评估待测个体是否为精神分裂症超高危人群的诊断装置,其特征在於,所述诊断装置包括:

(1) 检测单元,该检测单元包括检测待测个体DNA甲基化特征的试剂、和/或检测待测个体前脉冲抑制特征的设备 and 声音素材;

优选地,所述待测个体为未用药者;

优选地,所述DNA甲基化特征为权利要求3中所述的DNA甲基化特征;

优选地,所述前脉冲抑制特征为权利要求2中所述的前脉冲抑制特征;

(2) 分析单元,该分析单元用于对检测单元所得的检测结果进行分析,评估待测个体是否为精神分裂症超高危人群;

优选地,所述分析单元还包括采用随机森林算法建模得到包含DNA甲基化特征的DNA甲基化模型、和/或包含前脉冲抑制特征的前脉冲抑制模型;

更优选地,所述分析单元还包括对DNA甲基化模型、前脉冲抑制模型采用10折交叉验证得出最终模拟值,并基于所述最终模拟值构建包含DNA甲基化特征和前脉冲抑制特征的组合模型。

7. 一种用于评估待测个体精神分裂症患病风险的诊断装置,其特征在於,所述诊断装置包括:

(1) 检测单元,该检测单元包括检测待测个体DNA甲基化特征的试剂、和/或检测待测个体前脉冲抑制特征的设备 and 声音素材;

优选地,所述检测单元用于对首发精神分裂症患者、精神分裂症超高危人群、健康人群的DNA甲基化特征和前脉冲抑制特征进行检测;

优选地,所述待测个体为未用药者;

优选地,所述DNA甲基化特征为权利要求4中所述的DNA甲基化特征;

优选地,所述前脉冲抑制特征为权利要求2中所述的前脉冲抑制特征;

(2) 分析单元,该分析单元用于对检测单元所得的检测结果进行分析,评估待测个体精神分裂症的患病风险;

优选地,所述分析单元还包括采用随机森林算法建模得到包含DNA甲基化特征的DNA甲基化模型、和/或包含前脉冲抑制特征的前脉冲抑制模型;

更优选地,所述分析单元还包括对DNA甲基化模型、前脉冲抑制模型采用10折交叉验证得出最终模拟值,并基于所述最终模拟值构建包含DNA甲基化特征和前脉冲抑制特征的组合模型。

8. 一种用于精神分裂症早期检测和风险预测系统,其特征在于,所述系统包括如下模块:

(1) 测试模块,该测试模块用于对三类人群首发精神分裂症患者、精神分裂症超高危人群、健康人群的DNA甲基化特征和前脉冲抑制特征进行检测;

(2) 采集及预处理模块,该采集及预处理模块应用DNA甲基化芯片分析技术、前脉冲抑制测量技术,采集所述三类人群的DNA甲基化特征、和/或前脉冲抑制特征,并对所述DNA甲基化特征、和/或前脉冲抑制特征数据进行预处理,提取可应用于机器学习的数据;

(3) 数据分析模块,该数据分析模块对采集及预处理模块处理后的DNA甲基化特征数据、和/或前脉冲抑制特征数据进行缺失值处理、高偏差变量去除,归一化、标准化,特征降维、特征选择处理;

优选地,所述特征选择处理包括对DNA甲基化特征的选择处理;

更优选地,所述对DNA甲基化特征的选择处理包括使用R软件中的rfe函数通过多元递归特征消除方法进行特征的选择;

最优选地,所述rfe函数首先使用袋装树算法将模型拟合到所有特征,每个特征根据其模型的重要性进行排序,在特征选择的每次迭代中,保留已排序的特征,重新调整模型并评估性能,每组特征最终基于10折交叉验证进行选择;

(4) 学习分类模块,该学习分类模块使用数据分析模块处理分析得到的DNA甲基化特征数据、和/或前脉冲抑制特征数据训练模型,利用随机森林方法对所述三类人群进行分类,构建得到用于鉴别诊断所述三类人群的模型;

(5) 分类性能评估模块,该分类性能评估模块基于学习分类模块的结果,比较基于DNA甲基化特征、和/或前脉冲抑制特征构建得到的模型分类效果,计算正确率、敏感度、特异度、阳性预测值、阴性预测值、ROC曲线下面积AUC值对模型进行评估。

9. 根据权利要求8所述的系统,其特征在于,经分类性能评估模块分析评估得到的用于对首发精神分裂症患者、精神分裂症超高危人群、健康人群进行鉴别诊断的模型包括DNA甲基化特征、和/或前脉冲抑制特征;

优选地,所述DNA甲基化特征为权利要求4中所述的DNA甲基化特征;

优选地,所述前脉冲抑制特征为权利要求2中所述的前脉冲抑制特征。

10. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质包括存储的计算机程序,所述计算机程序被处理器执行时实现权利要求5所述的诊断装置、权利要求6所述的诊断装置、权利要求7所述的诊断装置、和/或权利要求8或9所述的系统。

DNA甲基化特征和前脉冲抑制特征在精神分裂症诊断中的应用

技术领域

[0001] 本发明属于生物医药技术领域,具体而言,本发明涉及DNA甲基化特征和前脉冲抑制特征在精神分裂症诊断中的应用。

背景技术

[0002] 精神分裂症(Schizophrenia, SCH)是一种以异常社会行为为特征的严重精神障碍疾病,也是一种病因尚未明确的复杂精神疾病。该病的主要临床表现有感知、思维、情感、行为等多方面的障碍以及精神活动的不协调,减少社会参与程度,缺乏动力等。精神分裂症患者还往往有额外的精神健康问题,如焦虑,抑郁状态等。严重影响患者的身体功能和生活质量。精神分裂症患者的相关症状通常从青春期开始逐渐出现并持续发展。流行病学调查显示精神分裂症在世界范围内的终身患病率为1%左右,据估算我国目前有700-800万精神分裂症患者,由此每年所造成的治疗及康复费用、以及对社会及家庭的负担是巨大的。

[0003] 绝大多数精神分裂症患者在首次发病前的一段时间内就已存在感知、思维、言语、行为等多方面的异常(也可称为“亚临床状态”),这段时间称为精神分裂症前驱期(Prodromal stage),此时期常见的症状包括:猜疑、奇特想法、抑郁、焦虑、情绪不稳、易激惹、记忆障碍、注意力不集中、对自我、他人及外界感知的变化以及睡眠障碍、躯体不适等。有前驱期表现的人群发展为精神分裂症的可能较大,这类人群目前被称为“超高危人群(Ultra high risk, UHR)”,精神分裂症超高危人群在1-2年的随访期内转化为精神分裂症的比例较高,其转化率可达9%-76%。对首发精神分裂症(First-episode schizophrenia, FES)患者而言,成功治疗后预防复发更为重要。

[0004] 精神分裂症的临床症状复杂多样,个体差异较大,虽然国际上已有一些较为严格的诊断标准如DSM-5、ICD-10等,国内也有常用的诊断标准如中国精神障碍分类与诊断标准CCMD-3,但是目前对精神分裂症的诊断方法主要依赖于医生根据病人的详细病史与精神症状,再参考其发病年龄、病期、病程等来做综合判断,由于医生的主观经验不同导致诊断结果也存在较大的差异性,因此目前对精神分裂症的诊断方法存在具有很强的主观性,缺乏客观的诊断指标等问题。近年来,为了使精神分裂症的诊断更加客观,减少人为因素,提高诊断的一致性和准确度,世界各地的科研工作者一直致力于寻找精神分裂症的生物标志物,建立有效的检测方法,以期提高精神分裂症诊断的一致性和准确度。

发明内容

[0005] 鉴于此,本发明的目的在于克服目前本领域诊断精神分裂症存在的技术缺陷,提高精神分裂症诊断的准确性和一致性,使精神分裂症的诊断更加客观,减少人为因素带来的主观误差,本发明提供了DNA甲基化特征和前脉冲抑制特征在精神分裂症诊断中的应用。

[0006] 本发明的上述目的通过以下技术方案得以实现:

[0007] 本发明的第一方面提供了DNA甲基化特征、和/或前脉冲抑制特征作为生物标志物

在制备诊断精神分裂症的产品中的应用。

[0008] 进一步,所述应用包括DNA甲基化特征、和/或前脉冲抑制特征作为生物标志物在制备诊断首发精神分裂症的产品中的应用;

[0009] 优选地,所述DNA甲基化特征包括900个DNA甲基化特征位点;

[0010] 更优选地,所述900个DNA甲基化特征位点包括cg25224760、cg25122934、cg25783352、cg00693060、cg23195763、cg25176297、cg18516940、cg22030766、cg21163478、cg15153114、cg03515346、cg16210324、cg18561199、cg24344143、cg03803284、cg23243856、cg27400655、cg13748287、cg14044124、cg12390678、cg18889956、cg07746811、cg25799064、cg15796941、cg24727290、cg10924691、cg11631427、cg26369259、cg22594586、cg01511844、cg15782771、cg18065318、cg25672411、cg23147317、cg18718234、cg15430422、cg22910549、cg06759901、cg25242729、cg22879458、cg25920734、cg25763426、cg20592447、cg04570624、cg20793712、cg25929589、cg04008557、cg10294363、cg23879228、cg12904004、cg04584468、cg24025893、cg17173330、cg01758041、cg17681524、cg24399801、cg07745707、cg25567990、cg10164186、cg21686797、cg17250082、cg13054419、cg25122125、cg13384409、cg27475680、cg26619156、cg24982343、cg17391741、cg11080430、cg24981400、cg19940644、cg20781380、cg20901993、cg13850825、cg02160684、cg24365551、cg22027671、cg16541340、cg15126273、cg21239409、cg18003231、cg24766555、cg25178683、cg21351992、cg11508828、cg24913623、cg16522885、cg18411015、cg22930950、cg17058296、cg27175294、cg19129200、cg12285326、cg11321371、cg24172989、cg20927425、cg10482356、cg04739228、cg14892768、cg22790758、cg02860199、cg11688731、cg20795889、cg06107159、cg17877898、cg16781264、cg06508379、cg15525341、cg27389290、cg27613174、cg26095809、cg10733080、cg14427074、cg13245883、cg24371156、cg16356539、cg01997461、cg17036570、cg23464565、cg27259382、cg24554865、cg26473110、cg10120112、cg23856267、cg05283486、cg25818522、cg24451839、cg21362233、cg09311571、cg19136704、cg20100445、cg17365824、cg22007638、cg14478036、cg27124109、cg01367037、cg23976937、cg00419512、cg13918430、cg14700479、cg14617167、cg26765871、cg13088598、cg02172133、cg19011911、cg17195231、cg26315559、cg11623950、cg26202181、cg21722655、cg06065611、cg26124670、cg27652200、cg01573019、cg17830740、cg19506491、cg26030037、cg17717125、cg22643214、cg17569754、cg11202345、cg20471798、cg26686249、cg18944640、cg22064182、cg17215601、cg27289295、cg02237470、cg26862091、cg13823585、cg09315290、cg23180925、cg24524946、cg19733221、cg24748771、cg22894329、cg13042714、cg23563927、cg08134068、cg25418474、cg20965876、cg25139963、cg20851097、cg10554947、cg27022602、cg25603927、cg12571055、cg10555383、cg04950413、cg07768486、cg26110834、cg21059366、cg24132991、cg19854991、cg26616258、cg23671319、cg04538470、cg07343706、cg23892310、cg22871908、cg02677877、cg24694879、cg01699425、cg24953213、cg27465618、cg25267487、cg10851774、cg02348488、cg18885125、cg25739742、cg00064261、cg09694403、cg18063493、cg15049549、cg23881601、cg08331138、cg23201527、cg26916621、cg22933195、cg22908033、cg04190450、cg24693635、cg17138984、cg16611213、cg09500174、cg09439920、cg15028904、cg24247231、cg06520003、cg00356723、cg27643917、cg25902889、cg15709265、cg22079361、cg14609682、cg21518713、cg11002709、cg17031944、cg01811230、cg21378920、

cg19362632、cg16260889、cg09319649、cg08057037、cg02100464、cg27272679、cg14880340、
cg24013954、cg24957704、cg22013228、cg18301955、cg10452175、cg05587552、cg23572751、
cg05156242、cg06884029、cg07108214、cg23847843、cg12858166、cg11868809、cg15282731、
cg05393023、cg02384255、cg03538137、cg12265130、cg13848607、cg16254946、cg27141176、
cg06045225、cg11568925、cg12062504、cg10926330、cg26930230、cg22332577、cg24836607、
cg02098075、cg26924679、cg14650228、cg11384205、cg00847359、cg13525067、cg16482175、
cg13036702、cg19120496、cg13857168、cg22413912、cg09358961、cg05328885、cg23062115、
cg20563131、cg18394235、cg09756436、cg06642177、cg19308620、cg20728314、cg01142417、
cg19802477、cg01072259、cg08722290、cg22910769、cg10117603、cg00551487、cg08314949、
cg09941712、cg06525127、cg01035812、cg17942689、cg19172429、cg05842490、cg18594106、
cg05612346、cg17389077、cg24964635、cg25621420、cg02203528、cg17654766、cg24957609、
cg20296990、cg17426070、cg14776168、cg04744810、cg22270828、cg00934864、cg04756223、
cg18116174、cg10286969、cg04641380、cg21778244、cg19251740、cg16619193、cg26837906、
cg19229071、cg19851029、cg01858657、cg06202930、cg10588705、cg04900856、cg10935762、
cg26220033、cg26259271、cg15560837、cg18793005、cg15663074、cg27159443、cg07134804、
cg19977871、cg26945392、cg08040115、cg04855546、cg18038737、cg17714225、cg08638514、
cg27173374、cg26643617、cg16993936、cg01894875、cg07349169、cg19988408、cg08205921、
cg00459299、cg05061471、cg14648186、cg11812015、cg14504291、cg21040513、cg16629408、
cg09370941、cg18328477、cg18266383、cg17876742、cg17730591、cg11775190、cg17343033、
cg07340025、cg07072638、cg10390228、cg14998917、cg19564077、cg12251779、cg03024403、
cg13719628、cg15137408、cg17335199、cg09650487、cg12298697、cg12468966、cg09603833、
cg04369787、cg16467919、cg21544091、cg11488893、cg21896553、cg02812534、cg18443412、
cg08501989、cg24257454、cg10880459、cg05665447、cg12331332、cg16937902、cg16316042、
cg04535320、cg13996738、cg00608943、cg13676583、cg18160999、cg21675569、cg01217923、
cg13651234、cg25197240、cg21548152、cg01772073、cg14382976、cg24241688、cg13426198、
cg01841782、cg17936572、cg16178855、cg00369126、cg15381380、cg13174660、cg15851317、
cg12401644、cg24365633、cg03828570、cg11590902、cg24052284、cg13692650、cg05064489、
cg24055461、cg13789353、cg00234363、cg16385933、cg09888026、cg07957070、cg18582505、
cg19132701、cg12808194、cg10121816、cg19691425、cg01926051、cg19324027、cg16470919、
cg12320972、cg02131013、cg08774990、cg06050385、cg10967622、cg15937259、cg18623931、
cg19484548、cg11002033、cg05682017、cg14429240、cg05197515、cg01438434、cg08877624、
cg22252041、cg05637384、cg17953869、cg10718940、cg22138494、cg13565624、cg06647133、
cg10220104、cg26291385、cg10249373、cg22044892、cg12483005、cg27229471、cg25370362、
cg02477579、cg06468646、cg02414586、cg02483627、cg13505944、cg09084892、cg07237023、
cg00232772、cg09200468、cg02581147、cg01428397、cg10157558、cg08464954、cg21403762、
cg06037240、cg05919900、cg00867317、cg02215945、cg14473016、cg11297423、cg04305319、
cg07322293、cg18793773、cg15600512、cg07870479、cg16278747、cg11605394、cg13096007、
cg08091867、cg24913501、cg03224013、cg00169917、cg20695587、cg02606535、cg16353907、
cg02968116、cg05976544、cg05236879、cg09721840、cg07599136、cg15887150、cg03005121、

cg08510930、cg01072942、cg19212949、cg09528372、cg01831459、cg15044205、cg08967338、
cg07848764、cg03644971、cg02894527、cg01360325、cg00700866、cg18504368、cg05141656、
cg02263704、cg00524443、cg23402560、cg09340485、cg18279180、cg17424473、cg15701419、
cg17509462、cg08845028、cg12290190、cg17084361、cg15405625、cg00723961、cg16183122、
cg06538872、cg03217733、cg22168617、cg07044458、cg21159260、cg08096851、cg03705571、
cg09099893、cg20226150、cg21455883、cg07938869、cg20336788、cg06814654、cg20643659、
cg05271612、cg00576340、cg12161959、cg11369706、cg04425202、cg11703939、cg03114004、
cg01705612、cg11706083、cg12388760、cg01771651、cg10701640、cg10691323、cg09414156、
cg01627046、cg17344560、cg03986909、cg09936757、cg01606576、cg05933932、cg15646610、
cg07066790、cg06458679、cg15565231、cg04734610、cg04526750、cg04290747、cg03204444、
cg14453769、cg14500267、cg13065504、cg03765047、cg12888632、cg00389386、cg14843586、
cg12402427、cg10703338、cg12394910、cg11036189、cg13515074、cg10440450、cg20355084、
cg10075524、cg08071164、cg07678448、cg08193100、cg05016746、cg05580671、cg05718343、
cg07431064、cg02843332、cg08741842、cg05497345、cg02540504、cg01582438、cg00369658、
cg04181189、cg02404489、cg00338113、cg00770663、cg00031362、cg00082739、cg00107488、
cg00203089、cg00223046、cg00230381、cg00321703、cg00417419、cg00459043、cg00480856、
cg00540722、cg00578363、cg00713465、cg00771653、cg00787055、cg00814990、cg00838379、
cg00934564、cg00948102、cg01102220、cg01148568、cg01158804、cg01261775、cg01372572、
cg01385063、cg01422467、cg01479738、cg01508386、cg01566396、cg01585372、cg01727923、
cg02039058、cg02068505、cg02173970、cg02278803、cg02292206、cg02447937、cg02481842、
cg02484210、cg02545393、cg02641277、cg02642822、cg02650286、cg02685016、cg02708898、
cg02724747、cg02741655、cg02794096、cg02800362、cg02803629、cg02873991、cg02916932、
cg02921122、cg03031823、cg03110787、cg03123782、cg03131097、cg03145322、cg03201507、
cg03223467、cg03303025、cg03447137、cg03505995、cg03550508、cg03553910、cg03600697、
cg03608224、cg03619332、cg03643709、cg03741350、cg03764027、cg03799192、cg03847642、
cg03849780、cg03864121、cg03880355、cg03892062、cg04001842、cg04059695、cg04101729、
cg04113225、cg04201335、cg04201365、cg04251616、cg04294058、cg04477962、cg04514292、
cg04555107、cg04619882、cg04791145、cg04832450、cg04833648、cg04843801、cg04889960、
cg04925085、cg04939919、cg04962756、cg04967200、cg05031519、cg05116088、cg05119778、
cg05221370、cg05269534、cg05282641、cg05333753、cg05341567、cg05369351、cg05399718、
cg05470179、cg05529874、cg05571437、cg05572751、cg05607935、cg05625951、cg05659199、
cg05899224、cg05949034、cg05971212、cg05977964、cg05978571、cg06100147、cg06185734、
cg06250386、cg06307176、cg06508886、cg06527318、cg06550214、cg06588466、cg06589239、
cg06613034、cg06627532、cg06760899、cg06807379、cg06813554、cg06909254、cg07008701、
cg07078225、cg07119973、cg07176514、cg07177889、cg07245558、cg07310500、cg07346747、
cg07372520、cg07403258、cg07618581、cg07747241、cg07754486、cg07850987、cg07936305、
cg07939245、cg07972159、cg08088948、cg08104023、cg08141194、cg08305707、cg08362804、
cg08371190、cg08578136、cg08611508、cg08616234、cg08791782、cg08817937、cg08864042、
cg08894540、cg08915378、cg08944086、cg08951822、cg08955275、cg09050775、cg09147024、

cg09163921、cg09173939、cg09250965、cg09262552、cg09382942、cg09444206、cg09460462、cg09465855、cg09514717、cg09524686、cg09528462、cg09615620、cg09619883、cg09631044、cg09678035、cg09682198、cg09694764、cg09726046、cg09747638、cg09802873、cg09806262、cg09983708、cg10033694、cg10052615、cg10158249、cg10163338、cg10169976、cg10214132、cg10247383、cg10343071、cg10488141、cg10488637、cg10549357、cg10587782、cg10589249、cg10730208、cg10758986、cg10803722、cg10841956、cg10844118、cg10973146、cg10989690、cg11021611、cg11029191、cg11046887、cg11051139、cg11187110、cg11251399、cg11466109、cg11466321、cg11585769、cg11611011、cg11625897、cg11673867、cg11709065、cg11770888、cg11836369、cg11854579、cg11873026、cg11995137、cg11999199、cg12005153、cg12092351、cg12099459、cg12105770、cg12107522、cg12154434、cg12192122、cg12228478、cg12332083、cg12470114、cg12472085、cg12500602、cg12507168、cg12508336、cg12553110、cg12595444、cg12666921、cg12832751、cg12841061、cg12884605、cg13013644、cg13248789、cg13254556、cg13282635、cg13318914、cg13413789、cg13435263、cg13443092、cg13455586、cg13456654、cg13471352、cg13476139、cg13621410、cg13629807、cg13647660、cg13708908、cg13714987、cg13728439、cg13731422、cg13749070、cg13779963、cg13807254、cg13962664、cg13962718、cg13994177、cg13999554、cg14019951、cg14036868、cg14122836、cg14125353、cg14131316、cg14221825、cg14280583;

[0011] 优选地,所述前脉冲抑制特征包括肌电变化百分比;

[0012] 更优选地,所述肌电变化百分比(%)的计算公式为:肌电变化百分比(%) = $(1 - S2/S1) \times 100\%$;

[0013] 其中,S1为仅含惊刺激的前脉冲抑制的肌电幅值,S2为同时含前刺激的前脉冲抑制的肌电幅值;

[0014] 最优选地,所述惊刺激为白噪声,持续时间40ms,声压级100dB;

[0015] 最优选地,所述前刺激为白噪声,分为左声道或右声道领先3ms,持续时间150ms,声压级65dB。

[0016] 在本发明的具体实施例中,所述DNA甲基化特征是基于Illumina Infinium MethylationEPIC BeadChip(850k)甲基化芯片采用重亚硫酸盐方法识别甲基化位点检测分析得到的DNA甲基化差异位点数据,进行进一步的处理分析得到的,所述处理包括对上述提取得到的数据进行缺失值处理、高偏差变量去除,归一化、标准化,特征降维、特征选择处理,提高模型的泛化能力,使用R软件中的rfe函数通过多元递归特征消除方法用于DNA甲基化的特征选择。rfe首先使用袋装树算法将模型拟合到所有特征。每个特征根据其对模型的重要性进行排序。在特征选择的每次迭代中,保留已排序的特征,重新调整模型并评估性能。每组特征最终基于10折交叉验证进行选择。

[0017] 在本发明的具体实施例中,所述前脉冲抑制特征是采用人类震惊反射系统(天鸣宏远人类震惊反射定制系统)在保持一定亮度和温度的听觉隔音屏蔽室中进行测量得到的,本发明中所用到的声音素材均由使用MATLAB函数库中的“randn()”生成,抽样率为48kHz,所述声音素材包括背景声、前刺激、惊刺激,统计的指标包括肌电变化百分比(%) = $(1 - S2/S1) \times 100\%$;肌电最大值潜伏期;肌电最大反应速度;持续注意水平。其中,刺激前500ms记为基线S0;仅含惊刺激记为S1,同时含前刺激记为S2;持续注意水平包括(但不限

于)持续盯住十字的时间。

[0018] 进一步,所述应用包括DNA甲基化特征、和/或前脉冲抑制特征作为生物标志物在制备诊断精神分裂症超高危人群的产品中的应用;

[0019] 优选地,所述DNA甲基化特征包括4个DNA甲基化特征位点;

[0020] 更优选地,所述4个DNA甲基化特征位点包括cg25376875、cg27415006、cg09543727、cg00213281;

[0021] 优选地,所述前脉冲抑制特征为前述前脉冲抑制特征。

[0022] 进一步,所述应用包括DNA甲基化特征、和/或前脉冲抑制特征作为生物标志物在制备鉴别诊断首发精神分裂症患者、精神分裂症超高危人群、健康人群的产品中的应用;

[0023] 优选地,所述DNA甲基化特征包括10个DNA甲基化特征位点;

[0024] 更优选地,所述10个DNA甲基化特征位点包括cg01511844、cg23076086、cg01807407、cg15782771、cg04538470、cg04218099、cg10297617、cg14240820、cg25929589、cg10315533;

[0025] 优选地,所述前脉冲抑制特征为前述前脉冲抑制特征。

[0026] 本发明的第二方面提供了一种用于评估待测个体患首发精神分裂症风险的诊断装置。

[0027] 进一步,所述诊断装置包括:

[0028] (1)检测单元,该检测单元包括检测待测个体DNA甲基化特征的试剂、和/或检测待测个体前脉冲抑制特征的设备 and 声音素材;

[0029] 优选地,所述待测个体为未用药者;

[0030] 优选地,所述DNA甲基化特征为本发明第一方面所述的900个DNA甲基化特征;

[0031] 优选地,所述前脉冲抑制特征为本发明第一方面所述的前脉冲抑制特征;

[0032] (2)分析单元,该分析单元用于对检测单元所得的检测结果进行分析,评估待测个体精神分裂症的患病风险;

[0033] 优选地,所述分析单元还包括采用随机森林算法建模得到包含DNA甲基化特征的DNA甲基化模型、和/或包含前脉冲抑制特征的前脉冲抑制模型;

[0034] 更优选地,所述分析单元还包括对DNA甲基化模型、前脉冲抑制模型采用10折交叉验证得出最终模拟值,并基于所述最终模拟值构建包含DNA甲基化特征和前脉冲抑制特征的组合模型。

[0035] 本发明的第三方面提供了一种用于评估待测个体是否为精神分裂症超高危人群的诊断装置。

[0036] 进一步,所述诊断装置包括:

[0037] (1)检测单元,该检测单元包括检测待测个体DNA甲基化特征的试剂、和/或检测待测个体前脉冲抑制特征的设备 and 声音素材;

[0038] 优选地,所述待测个体为未用药者;

[0039] 优选地,所述DNA甲基化特征为本发明第一方面所述的4个DNA甲基化特征;

[0040] 优选地,所述前脉冲抑制特征为本发明第一方面所述的前脉冲抑制特征;

[0041] (2)分析单元,该分析单元用于对检测单元所得的检测结果进行分析,评估待测个体精神分裂症的患病风险;

[0042] 优选地,所述分析单元还包括采用随机森林算法建模得到包含DNA甲基化特征的DNA甲基化模型、和/或包含前脉冲抑制特征的前脉冲抑制模型;

[0043] 更优选地,所述分析单元还包括对DNA甲基化模型、前脉冲抑制模型采用10折交叉验证得出最终模拟值,并基于所述最终模拟值构建包含DNA甲基化特征和前脉冲抑制特征的组合模型。

[0044] 本发明的第四方面提供了一种用于评估待测个体精神分裂症患病风险的诊断装置。

[0045] 进一步,所述诊断装置包括:

[0046] (1) 检测单元,该检测单元包括检测待测个体DNA甲基化特征的试剂、和/或检测待测个体前脉冲抑制特征的设备和声音素材;

[0047] 优选地,所述检测单元用于对首发精神分裂症患者、精神分裂症超高危人群、健康人群的DNA甲基化特征和前脉冲抑制特征进行检测;

[0048] 优选地,所述待测个体为未用药者;

[0049] 优选地,所述DNA甲基化特征为本发明第一方面所述的10个DNA甲基化特征;

[0050] 优选地,所述前脉冲抑制特征为本发明第一方面所述的前脉冲抑制特征;

[0051] (2) 分析单元,该分析单元用于对检测单元所得的检测结果进行分析,评估待测个体精神分裂症的患病风险;

[0052] 优选地,所述分析单元还包括采用随机森林算法建模得到包含DNA甲基化特征的DNA甲基化模型、和/或包含前脉冲抑制特征的前脉冲抑制模型;

[0053] 更优选地,所述分析单元还包括对DNA甲基化模型、前脉冲抑制模型采用10折交叉验证得出最终模拟值,并基于所述最终模拟值构建包含DNA甲基化特征和前脉冲抑制特征的组合模型。

[0054] 本发明的第五方面提供了一种用于精神分裂症早期检测和风险预测系统。

[0055] 进一步,所述系统包括如下模块:

[0056] (1) 测试模块,该测试模块用于对三类人群首发精神分裂症患者、精神分裂症超高危人群、健康人群的DNA甲基化特征和前脉冲抑制特征进行检测;

[0057] (2) 采集及预处理模块,该采集及预处理模块应用DNA甲基化芯片分析技术、前脉冲抑制测量技术,采集所述三类人群的DNA甲基化特征、和/或前脉冲抑制特征,并对所述DNA甲基化特征、和/或前脉冲抑制特征数据进行预处理,提取可应用于机器学习的数据;

[0058] (3) 数据分析模块,该数据分析模块对采集及预处理模块处理后的DNA甲基化特征数据、和/或前脉冲抑制特征数据进行缺失值处理、高偏差变量去除,归一化、标准化,特征降维、特征选择处理;

[0059] 优选地,所述特征选择处理包括对DNA甲基化特征的选择处理;

[0060] 更优选地,所述对DNA甲基化特征的选择处理包括使用R软件中的rfe函数通过多元递归特征消除方法进行特征的选择;

[0061] 最优选地,所述rfe函数首先使用袋装树算法将模型拟合到所有特征,每个特征根据其模型的重要性进行排序,在特征选择的每次迭代中,保留已排序的特征,重新调整模型并评估性能,每组特征最终基于10折交叉验证进行选择;

[0062] (4) 学习分类模块,该学习分类模块使用数据分析模块处理分析得到的DNA甲基化

特征数据、和/或前脉冲抑制特征数据训练模型,利用随机森林方法对所述三类人群进行分类,构建得到用于鉴别诊断所述三类人群的模型;

[0063] (5)分类性能评估模块,该分类性能评估模块基于学习分类模块的结果,比较基于DNA甲基化特征、和/或前脉冲抑制特征构建得到的模型分类效果,计算正确率、敏感度、特异度、阳性预测值、阴性预测值、ROC曲线下面积AUC值对模型进行评估。

[0064] 进一步,经分类性能评估模块分析评估得到的用于对首发精神分裂症患者、精神分裂症超高危人群、健康人群进行鉴别诊断的模型包括DNA甲基化特征、和/或前脉冲抑制特征;

[0065] 优选地,所述DNA甲基化特征为本发明第一方面中所述的10个DNA甲基化特征;

[0066] 优选地,所述前脉冲抑制特征为本发明第一方面中所述的前脉冲抑制特征。

[0067] 本发明的第六方面提供了一种计算机可读存储介质。

[0068] 进一步,所述计算机可读存储介质包括存储的计算机程序,所述计算机程序被处理器执行时实现本发明第二方面所述的诊断装置、本发明第三方面所述的诊断装置、本发明第四方面所述的诊断装置、和/或本发明第五方面所述的系统。

[0069] 精神分裂症由于其病因的复杂性及症状的异质性,本领域一直缺乏可用于临床实践的客观生物学指标,在本发明中,DNA甲基化作为表观遗传调控的常见形式之一,是一个基于基因同时又受环境事件影响的指标;而前脉冲抑制(Prepulse inhibition,PPI)作为精神分裂症的内表型之一,与疾病的典型行为症状密切相关,同时携带有疾病的遗传负荷。本发明通过分析筛选验证得到包含特定DNA甲基化特征、和/或前脉冲抑制特征的诊断模型,可用于精神分裂症诊断和鉴别,且具有较好的诊断效能。

[0070] 相对于现有技术,本发明具有的优点和有益效果:

[0071] 本发明首次将DNA甲基化特征和前脉冲抑制特征用作临床生物标志物,联合应用于诊断鉴别首发精神分裂症、精神分裂症超高危人群和健康人群,具有客观、准确、有效等优点,在临床上具有非常好的应用前景,为精神分裂症的早期识别和干预奠定了基础,为精准医学策略的开发提供了支持。

附图说明

[0072] 以下,结合附图来详细说明本发明的实施方案,其中:

[0073] 图1显示本发明的技术路线图;

[0074] 图2显示本发明所述前脉冲抑制的测试范式示意图;

[0075] 图3显示DNA甲基化模型、PPI模型、组合特征(DNA甲基化特征和PPI特征联合)模型对首发精神分裂症诊断的ROC曲线图,其中,横坐标FPR,Fales Positive Rate,假正率;纵坐标TPR:True Positive Rate,真正率;

[0076] 图4显示首发精神分裂症区分健康人群概率分布图,其中,横坐标表示DNA甲基化模型模拟值,数值范围自0-100%,数值越大代表个体为首发精神分裂症患者概率越高;纵坐标表示PPI模型模拟值,数值范围自0-100%,数值越大代表个体为首发精神分裂症患者概率越高;三角形CON,表示个体真实为健康人群;圆形FES,表示个体真实为首发精神分裂症患者;

[0077] 图5显示DNA甲基化模型、PPI模型、组合特征(DNA甲基化特征和PPI特征联合)模型

对超高危人群诊断的ROC曲线图,其中,横坐标FPR,Fales Positive Rate,假正率;纵坐标TPR:True Positive Rate,真正率;

[0078] 图6显示超高危人群区分健康人群概率分布图,其中,横坐标表示DNA甲基化模型模拟值,数值范围自0-100%,数值越大代表个体为超高危人群概率越高;纵坐标表示PPI模型模拟值,数值范围自0-100%,数值越大代表个体为超高危人群概率越高;三角形CON,表示个体真实为健康人群;圆形UHR,表示个体真实为超高危人群;

[0079] 图7显示基于DNA甲基化模型的三组人群UMAP (Uniform Manifold Approximation and Projection) 图,其中,横纵坐标为经UMAP特征降维算法计算的DNA甲基化特征二维主成分;三角形表示CON,健康对照人群;正方形表示UHR,超高危人群;圆形表示FES,首发精神分裂症患者。

具体实施方式

[0080] 下面结合具体实施例,进一步阐述本发明,仅用于解释本发明,而不能理解为对本发明的限制。本领域的普通技术人员可以理解为:在不脱离本发明的原理和宗旨的情况下可以对这些实施例进行多种变化、修改、替换和变型,本发明的范围由权利要求及其等同物限定。下列实施例中未注明具体条件的实验方法,通常按照常规条件或按照厂商所建议的条件实施检测。

[0081] 实施例精神分裂症诊断模型的构建及其诊断效能的验证

[0082] 在本研究中,共纳入31例首发精神分裂症患者、34例精神分裂症超高危人群、26例健康对照。通过Illumina MethylationEPIC甲基化芯片收集受试者DNA甲基化数据,通过知觉空间分离范式收集PPI数据。对DNA甲基化数据进行归一化、标准化之后,利用递归特征消除法筛选特征。分别以DNA甲基化特征、前脉冲抑制特征以及这两组的特征组合构建模型,应用随机森林算法通过10折交叉验证区分三组人群,具体技术路线图见图1。

[0083] 1、研究对象

[0084] 本研究共招募3组受试者:首次发病未用药的精神分裂症患者(First-episode schizophrenia, FES) 31例、超高危人群(Ultra high risk, UHR) 34例、健康对照人群(Health control, CON) 各26例。首发精神分裂症患者组、超高危人群均是从首都医科大学附属北京安定医院门诊或住院部门招募;健康对照人群从工作人员、在读学生及社会上招募。本研究相关伦理问题都已通过首都医科大学附属北京安定医院伦理委员会审查。所有受试者皆告知参与研究的获益及风险,并签署知情同意书。

[0085] 入组标准:共用部分:自愿签署知情同意书并告知研究方案;右利手,年龄在18-55岁之间;初中及以上文化程度,智力测试正常($IQ \geq 75$);听力正常,既往无听力系统疾病;无酒精和药物滥用史;无神经系统疾病家族史(脑血管意外除外);目前无明确诊断的神经系统和其他严重躯体疾病病史。FES:符合DSM-5精神分裂症诊断标准,病程 < 3 年且未经过系统抗精神病药治疗者。UHR:符合SIPS筛查前驱期风险综合征标准。CON:符合DSM-5未患有任

[0086] 排除标准:体格检查和实验室生化指标或脑电图、心电图严重异常者;妊娠或哺乳期女性,随访期间有可能怀孕的女性;因其他躯体疾病不能中断使用药物者;3个月内参加过其他科研治疗、献血或作为受试者被采血者;极度兴奋、冲动、不合作者;研究者认为受试

者在研究过程中会有自杀或暴力行为者;6个月内接受过电抽搐或磁刺激治疗者。FES:目前无明确判断的除精神分裂症外的其他精神疾病;当前或曾经服用过抗精神病药、抗抑郁药、中枢兴奋剂、情感稳定剂治疗者。UHR、CON:现在或既往被诊断患有精神障碍者;现在或既往被诊断为脑器质性或躯体疾病所致或精神活性物质所致精神障碍;当前或曾服用过抗精神病药、抗抑郁药、中枢兴奋剂、情感稳定剂治疗者。

[0087] 2、实验方法

[0088] (1) 一般人口学资料及临床资料收集

[0089] 通过自制问卷的形式,由主试询问受试者的一般人口学资料,包括受试者年龄、性别、种族、文化程度、利手、烟酒情况、个人史、家族史等;如受试者为患者还将询问临床资料(如PANSS量表、发病年龄、用药治疗情况等);如受试者为超高危人群,还将进行精神病风险综合征定式访谈(The Structured Interview for Psychosis-risk Syndromes,SIPS)。由2名经严格培训并实施一致性评定合格的临床医师施测。

[0090] (2) 认知功能评估

[0091] 测试的认知维度包括:信息处理速度、注意/警觉、工作记忆、词语学习、视觉学习、推理和问题解决、社会认知。由3名经严格培训的主试进行施测。为避免被试疲劳,中间可有5~10分钟的休息。

[0092] (3) 前脉冲抑制测量

[0093] ①实验设备:天鸣宏远人类震惊反射定制系统。

[0094] ②实验环境:听觉隔音屏蔽室中进行,保持一定光亮和温度。

[0095] ③声音素材:所需声音素材皆由使用MATLAB函数库中的“randn()”生成,抽样率为48kHz。生成后的高斯噪声通过512阶、截止频率为10kHz的低通数字滤波器得到宽带噪声,作为背景噪声。另外生成长度为150ms宽带噪声的前刺激声音,及长度为40ms宽带噪声的惊刺激声音。为避免能量飞溅现象,每个声音刺激都是在前脉冲刺激声音及惊跳刺激声音引入后再滤波。声音信号利用声卡(Audio CODEC'97)输入到Sennheiser HD 600耳机呈现给被试。利用声压校正仪(Larson Davis,AUDit and System 824)进行声压校正。以下是各声音具体参数,a)背景声:白噪声,分为左声道或右声道领先3ms,持续时间15s,声压级60dB;b)前刺激:白噪声,分为左声道或右声道领先3ms,持续时间150ms,声压级65dB;c)惊刺激:白噪声,持续时间40ms,100dB。

[0096] ④测试流程:嘱被试坐在检查椅上,于右眼瞳孔下、外侧1.5cm贴2个Ag/AgCl电极记录眼轮匝肌电活动,于右乳突贴电极作为地线。眼睛距眼动仪显示器60cm。全身肌肉放松,嘱实验过程中尽量保持清醒,双眼注视屏幕,尽量保持头部固定不动。首先是对各声音素材进行学习,每个声音素材学习3遍,对所有被试给予相同的指导语,背景声和前刺激需被试分辨左声道或右声道领先,对被试进行单眼3点刻度眼动校准,嘱被试注视前方十字,开始测试。

[0097] ⑤测试范式:前刺激持续时间为150ms,分为左领先及右领先。前刺激与背景声(左领先和右领先)呈现出2种知觉空间分离状态(前刺激左领先,背景声右领先;前刺激右领先,背景声左领先),每种状态重复5次。前刺激和惊刺激间隔120ms。仅有惊刺激的试次为7次。具体可见图2。

[0098] ⑥统计指标:刺激前500ms记为基线S0;仅含惊刺激记为S1,同时含前刺激记为S2;

计算肌电变化百分比 (%) = $(1 - S_2/S_1) \times 100\%$; 肌电最大值潜伏期; 肌电最大反应速度; 持续注意水平 (如持续盯住十字的时间)。

[0099] (4) Illumina Infinium MethylationEPIC BeadChip (850k) 甲基化芯片检测

[0100] Illumina 850k DNA甲基化芯片 (所述芯片购自于Illumina公司) 采用重亚硫酸盐方法识别甲基化位点, 到达单碱基分辨率, 可检测人全基因组约853, 307个CpG位点的甲基化状态, 不但保持了对CpG岛, 基因启动子区的全面覆盖, 还特别加强了增强子区 (新增了333, 265个探针覆盖来自ENCODE及FANTOM5计划的增强子) 以及基因编码区的探针覆盖。具体实验流程如下:

[0101] ①样品质检: 提取DNA。对于基因组DNA, 先用分光光度计定量, 并将样品调到标准浓度50ng/ μ L, 20 μ L, 然后用0.8%的琼脂糖凝胶电泳。样品电泳主带清晰, 通常不小于10kb, 没有明显降解, 总量5 μ L以上, 可进行下游的甲基化芯片实验。

[0102] ②亚硫酸盐转化: 根据Illumina官方推荐的Zymo EZ DNA Methylation Kit优化方法进行亚硫酸盐转化。

[0103] ③DNA扩增: 制备MSA3板, 在样本中加入0.1N NaOH使DNA变性为单链, 经中和后加入全基因组扩增试剂, 在37 $^{\circ}$ C恒温条件下过夜孵育。

[0104] ④DNA的片段化、沉淀、重悬: DNA片段化: 扩增后产物, 经过酶解处理, 得到片段化DNA。DNA沉淀: 加入异丙醇在4 $^{\circ}$ C下离心沉淀DNA片段。DNA重悬: 沉淀后的DNA在空气中干燥后, 加入杂交缓冲试剂使DNA沉淀重新溶解。

[0105] ⑤DNA与芯片的杂交: 将重悬后的DNA样本与准备好的芯片杂交, 置于杂交炉内过夜。在杂交过程中, 片段化后的DNA经过变性, 与特异位点的50个碱基 (连接在芯片的微珠 (bead) 上) 退火。

[0106] ⑥芯片清洗、单碱基延伸、染色: 洗去未杂交和非特异杂交的DNA, 以捕获到的DNA为模板, 在芯片上进行单碱基的延伸反应, 在芯片上加上可检测的荧光基团, 从而区分样本的甲基化状况。包被芯片: 将反应完成的芯片放入XC4试剂中, 使其表面包裹上一层粘性透明液体, 再将其放入真空环境下干燥1小时, 从而将芯片包被, 保护其信号稳定较长的时间。

[0107] ⑦芯片扫描和数据提取: 提前下载相应的manifest文件, 将处理好的芯片放入扫描仪, 利用激光激发芯片上单碱基延伸产物的荧光基团, 扫描仪获取由荧光基团发出的荧光, 并生成原始数据, 记录扫描结果存放的位置。由此所得的数据直接导入R包ChAMP软件进行分析, 从而就得到每个样本每个位点的甲基化水平数据。

[0108] (5) DNA甲基化差异位点分析

[0109] 通过R软件 (R version 3.6.3) ChAMP包导入甲基化矩阵, 获得每个位点的原始信号值和Detection P等信息。具体实验流程如下:

[0110] ①数据质控: 分为位点质控和个体质控。位点质控要求位点在95%以上的个体中的Detection P小于0.05且beadcount不小于3, 同时去除位于X, Y染色体及SNP上的位点; 个体质控要求个体在95%以上位点中的Detection P小于0.05。其次, 在批次校正之前, 本研究以“slide”和“position”作为批次变量, 移除了代表小批次 (≤ 5 个样本) 或不平衡批次的样本。

[0111] ②归一化处理: 在Illumina DNA甲基化微阵列上, 探针采用两种不同的设计 (称为I型探针和II型探针), 具有不同的杂交化学性质, 这意味着来自这两种不同设计的探针将

表现出不同的分布,并且这与由I型和II型探针的生物学特性(例如CpG密度)差异无关。I型探针和II型探针甲基化分布之间最显著的差异是II型探针分布表现出动态范围降低。基于数据质控过滤后的数据,使用BMIQ(beta-mixture quantile normalization)对探针类型偏倚(probe type bias)进行校正,得到最终可用于差异分析的甲基化水平(beta值)。

[0112] ③批次效应、细胞类型及其他混杂因素校正:具体过程如下,

[0113] a通过sva包ComBat算法,使用经验贝叶斯方法来校正技术差异校正批次效应。本研究以“slide”、“position”为批次变量。在使用ComBat调整之前通过log转换beta值,然后在ComBat调整之后重新转换为beta值。

[0114] b由于血液是由多种细胞类型的构成,同时有鉴于DNA甲基化具有高度的细胞类型特异性,因此分析不受细胞类型组成潜在变化驱动的差异甲基化位点还通常需要校正细胞类型异质性。RefbaseEWAS方法通过参考血液主要细胞类型的DNA甲基化谱数据库,以受约束的多元回归形式推断特定于样品的细胞类型比例,并以此作为协变量,计算不受细胞类型组成影响驱动的表达相关变化,校正甲基化矩阵。

[0115] c对于吸烟状况,本研究按照文献Elliott H R,Tillin T,Mcardle W L,et al.Differences in smoking associated DNA methylation patterns in South Asians and Europeans[J].Clinical epigenetics,2014,6(1):1-10中的方法,基于已知与吸烟状况相关的位点的DNA甲基化谱推算出DNA甲基化吸烟评分。这种方法基于183个DNA甲基化位点上产生一个加权得分,其中权重取自以往研究结果(Zeilinger S,Kühnel B,Klopp N,et al.Tobacco smoking leads to extensive genome-wide changes in DNA methylation [J].PloS one,2013,8(5):e63812.)。最后将年龄、性别、吸烟评分作为混杂因素进行校正。

[0116] ④差异甲基化位点(DMP)分析:利用R软件Limma包进行差异甲基化位点分析,Limma采用moderated t-statistics和empirical Bayes方法检验差异位点的显著性,同时针对多重假设检验问题计算校正后的P值。显著差异CpG位点默认筛选标准为:adjust.P.Value<0.05。

[0117] (6)基于DNA甲基化和PPI的机器学习模型区分研究对象

[0118] ①提取可应用于机器学习的数据,PPI评估、DNA甲基化资料。

[0119] ②特征选择:为提高分类准确率,对提取数据进行缺失值处理、高偏差变量去除,归一化、标准化,特征降维、特征选择处理,提高模型的泛化能力,使用R软件中的rfe函数通过多元递归特征消除方法用于DNA甲基化的特征选择。rfe首先使用袋装树算法将模型拟合到所有特征。每个特征根据其对模型的重要性进行排序。在特征选择的每次迭代中,保留已排序的特征,重新调整模型并评估性能。每组特征的最终基于10折交叉验证进行选择。

[0120] ③构建模型:在完成特征选择等数据处理后,使用提取的数据训练模型,使用的方法包括随机森林(random forest,RF)方法,本研究将使用R软件中的randomForest等程序包来构建精神分裂症的辅助诊断模型。

[0121] ④模型评估:使用正确率、敏感度、特异度、阳性预测值、阴性预测值、ROC曲线下面积对模型进行评估。

[0122] 3、实验结果

[0123] (1)人口学及临床特征

[0124] 本阶段研究共收集91例受试者,包括31例首发未用药精神分裂症患者、34例超高

危人群、26例健康对照。三组人群在年龄、性别、受教育年限上均不存在显著差异。在吸烟评分上,两两比较结果显示健康对照与其他两组均存在显著差异(Bonferroni校正后 $P < 0.05$),后续进一步DNA甲基化分析中对此变量进行了校正(见表1);

[0125] 在认知功能评估方面,健康对照与首发精神分裂症患者相比在认知功能各维度差异均有统计学意义,与超高危人群相比在工作记忆、认知总分上差异存在统计学意义。首发精神分裂症患者与超高危人群,在注意警觉、工作记忆、视觉学习、推理和问题解决、认知总分上差异存在统计学意义(见表1);

[0126] 三组受试者在知觉空间分离PPI (PSS-PPI) 上,健康对照与首发精神分裂症患者、超高危人群之间的差异均有统计学意义,首发精神分裂症患者与超高危人群间未见显著差异(见表1)。

[0127] 表1首次发病精神分裂症患者、超高危人群及健康受试者的一般人口学特征、MCCB认知功能评估、PPI评估结果

[0128]	项目	FES (n=31)	UHR (n=34)	CON (n=26)	P value	各组两两比较 (P, Bonferroni 法校正)		
						FESvsCON	FESvsUHR	UHRvsCON
	Gender (Male, %)	15 (48.4)	18 (52.9)	18 (69.2)	0.258 ^a	0.187 ^a	0.906 ^a	0.312 ^a
	Age (mean (SD))	23.97 (7.89)	23.35 (5.99)	23.85 (3.28)	0.913	0.938	0.727	0.686
	Onset (mean (SD))	21.68 (7.44)	NA	NA	NA	NA	NA	NA
	Edu (mean (SD))	12.45 (3.05)	13.68 (3.47)	13.00 (3.35)	0.329	0.524	0.135	0.448
	Smoking (mean (SD))	-7.00 (2.31)	-7.71 (2.04)	-9.92 (2.10)	<0.001	<0.001	0.593	<0.001
	SOP (mean (SD))	33.65 (7.88)	39.49 (7.53)	42.65 (7.04)	<0.001	<0.001	0.003	0.100
	AV (mean (SD))	31.82 (9.46)	40.37 (12.20)	43.58 (8.99)	<0.001	<0.001	0.002	0.246
	WM (mean (SD))	37.13 (9.31)	40.05 (10.58)	45.85 (6.83)	0.002	<0.001	0.241	0.013
[0129]	VBL (mean (SD))	37.53 (8.65)	43.24 (9.36)	47.46 (10.43)	0.001	<0.001	0.013	0.110
	VSL (mean (SD))	38.90 (14.72)	43.32 (11.85)	47.19 (8.86)	0.042	0.012	0.190	0.153
	RPS (mean (SD))	35.20 (11.46)	40.82 (10.38)	40.12 (10.20)	0.084	0.092	0.043	0.792
	SC (mean (SD))	34.12 (12.09)	37.17 (8.44)	39.82 (8.79)	0.101	0.045	0.247	0.246
	OCV (mean (SD))	35.81 (5.74)	40.85 (5.66)	43.72 (5.26)	<0.001	<0.001	0.001	0.047
	PANSS (mean (SD))	84.23 (12.84)	NA	NA	NA	NA	NA	NA
	P scores (mean (SD))	23.52 (5.33)	NA	NA	NA	NA	NA	NA
	N scores (mean (SD))	18.90 (7.42)	NA	NA	NA	NA	NA	NA
	G scores (mean (SD))	41.81 (7.09)	NA	NA	NA	NA	NA	NA
	PSS_PPI (mean (SD))	24.96 (23.42)	25.62 (15.68)	41.09 (15.50)	0.002	0.003	0.894	<0.001

[0130] FES:首发未用药精神分裂症患者;UHR:超高危人群;CON:健康对照;P value:方差分析或卡方检验;a.卡方检验;Gender:性别;Age:年龄;Onset:发病年龄;Edu:受教育年限;Smoking:吸烟评分;SOP:信息处理速度;AV:注意警觉;WM:工作记忆;VBL:词语学习;VSL:视觉学习;RPS:推理和问题解决;SC:社会认知;OCV:认知总分;PANSS:阳性和阴性症状量表;P Score:阳性量表得分;N Score:阴性量表得分;G Score:一般精神病理量表得分;PSS_PPI:知觉空间分离PPI;NA:不适用

[0131] (2) 基于DNA甲基化和PPI的机器学习模型区分首发患者与健康人群

[0132] 基于DNA甲基化和PPI的机器学习模型采用随机森林算法构建。将FES组标记为1, CON组标记为0,分别基于DNA甲基化和PPI,利用随机森林算法建模,其中DNA甲基化数据以校正后 $P < 0.05$ 为阈值,共筛选出18696差异甲基化位点,经过特征处理共筛得900个特征位点(900个DNA甲基化特征位点在发明内容部分进行了详细的列举)用于建模。再分别对DNA甲基化模型和PPI模型采用10折交叉验证(10-fold Cross Validation, 10-fold CV)得出

最终模拟值,同时基于DNA甲基化模拟值和PPI模拟值构建特征组合模型。最后针对DNA甲基化模型、PPI模型和特征组合模型分别绘制ROC曲线(Receiver Operating Characteristic curves,ROC curves),并以DNA甲基化模拟值为横坐标、PPI模拟值为纵坐标绘制首发精神分裂症区分健康人群概率分布图(见图3和4)。结果显示,DNA甲基化模型、PPI模型、组合特征模型准确率分别为82.5%、93.0%、94.7%,ROC曲线下面积(ROC Area Under Curvrs,AUC)分别为90.3%、98.8%、98.6%。总体上来说,DNA甲基化模型区分FES组与CON组表现不如PPI模型和组合特征模型,由DNA甲基化特征和PPI特征联合构成的组合特征模型对首发精神分裂症的诊断效能最佳(见表2)。

[0133] (3) 基于DNA甲基化和PPI的机器学习模型区分超高危人群与健康人群

[0134] 将UHR组标记为1,CON组标记为0,分别基于DNA甲基化和PPI,利用随机森林算法建模,其中DNA甲基化数据以校正后 $P < 0.05$ 为阈值,共筛选出47376差异甲基化位点,经过特征处理共筛得4个特征位点(4个DNA甲基化特征位点分别为cg25376875、cg27415006、cg09543727、cg00213281)用于建模。再分别对DNA甲基化模型和PPI模型采用10折交叉验证(10-fold Cross Validation,10-fold CV)得出最终模拟值,同时基于DNA甲基化模拟值和PPI模拟值构建特征组合模型。最后针对DNA甲基化模型、PPI模型和特征组合模型分别绘制ROC曲线(Receiver Operating Characteristic curves,ROC curves),并以DNA甲基化模拟值为横坐标、PPI模拟值为纵坐标绘制超高危人群区分健康人群概率分布图(见图5和图6)。结果显示,DNA甲基化模型、PPI模型、组合特征模型准确率分别为88.3%、90.0%、96.6%,ROC曲线下面积分别为(ROC Area Under Curvrs,AUC)96.7%、98.3%、99.0%。总体上来说,DNA甲基化模型区分UHR组与CON组表现稍弱于PPI模型和组合特征模型,组合特征模型对首超高危人群的诊断效能最佳(见表2)。

[0135] (4) 基于DNA甲基化和PPI的机器学习模型区分三组人群

[0136] 针对FES组、UHR组、CON组三组人群,分别基于DNA甲基化和PPI,利用随机森林算法建模,其中DNA甲基化数据以校正后 $P < 0.05$ 为阈值,共筛选出28245差异甲基化位点,经过特征处理共筛得10个特征位点(10个DNA甲基化特征位点分别为cg01511844、cg23076086、cg01807407、cg15782771、cg04538470、cg04218099、cg10297617、cg14240820、cg25929589、cg10315533)用于建模。再分别对DNA甲基化模型和PPI模型采用10折交叉验证(10-fold Cross Validation,10-fold CV)得出最终模拟值,同时基于DNA甲基化模拟值和PPI模拟值构建特征组合模型。利用一致流形逼近与投影(Uniform Manifold Approximation and Projection,UMAP)算法绘制基于DNA甲基化模型的三组人群UMAP图(见图7)。结果显示,DNA甲基化模型、PPI模型、组合特征模型准确率分别为79.1%、44.0%、79.1%。总体上来说,PPI模型表现远逊于DNA甲基化模型和组合特征模型,DNA甲基化模型和组合特征模型对三组人群(首发精神分裂症患者、超高危人群、健康人群)具有相对较好的区分效果(见表3)。

[0137] 表2基于DNA甲基化和PPI的机器学习模型区分首发精神分裂症患者与健康人群、超高危人群与健康人群

	准确率	敏感度	特异度	阳性预测值	阴性预测值	AUC (95% CI)
FES vs CON						
(%)						
DNA 甲基化特征	82.5	83.9	80.8	83.9	80.8	90.3 (82.4-98.2)
PPI 特征	93.0	90.3	96.2	96.6	89.3	98.8 (96.9-100)
特征组合	94.7	96.8	92.3	93.8	96.0	98.6 (96.4-100)
[0138]						
UHR vs CON						
(%)						
DNA 甲基化特征	88.3	88.2	88.5	90.9	85.2	96.7 (93.1-100)
PPI 特征	90.0	88.2	92.3	96.8	85.7	98.3 (96.1-100)
特征组合	96.6	97.1	96.2	97.1	96.2	99.0 (96.9-100)

[0139] 各模型模拟值,皆以0.5为划界值计算准确率、敏感度、特异度、阳性预测值、阴性预测值。FES vs CON表示模型区分首发精神分裂症患者与健康人群;UHR vs CON表示模型区分超高危人群与健康人群

[0140] 表3基于DNA甲基化和PPI的机器学习模型区分三组人群

[0141]	预测值	CON (n=26)	UHR (n=34)	FES (n=31)	准确率 (%)
--------	-----	------------	------------	------------	---------

		真实值			
	DNA 甲基化模型				79.1
	CON	17	3	2	
	UHR	5	28	2	
	FES	4	3	27	
	DNA 甲基化模型				44.0
[0142]	CON	12	9	7	
	UHR	8	17	13	
	FES	6	8	11	
	组合模型				79.1
	CON	17	3	2	
	UHR	5	28	2	
	FES	4	3	27	

[0143] 纵列代表CON、UHR、FES三组人群的真实个数,横行代表CON、UHR、FES经过机器学习模型预测后区分至三组人群中个数

[0144] 上述实施例的说明只是用于理解本发明的方法及其核心思想。应当指出,对于本领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以对本发明进行若干改进和修饰,这些改进和修饰也将落入本发明权利要求的保护范围内。

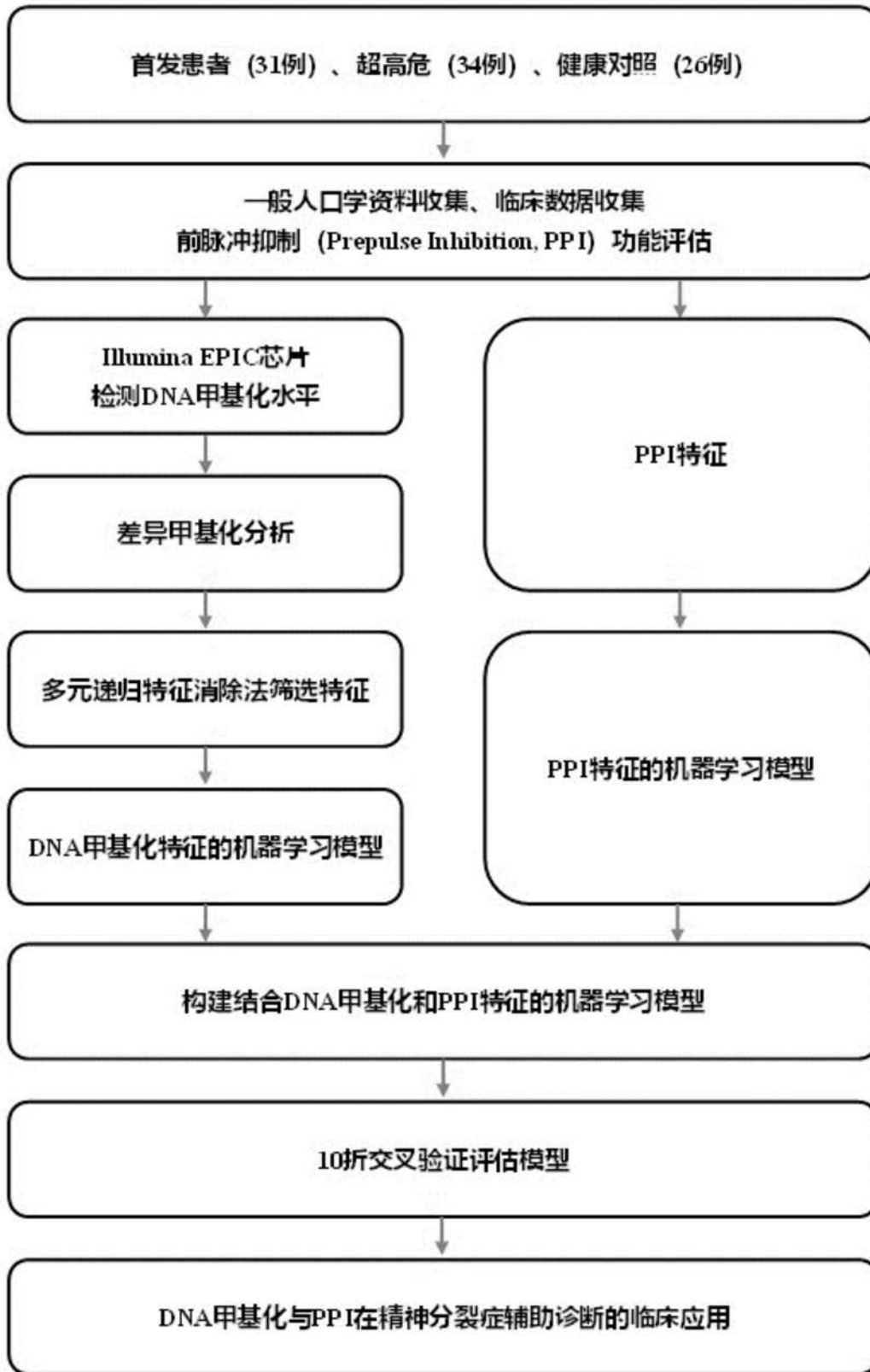


图1

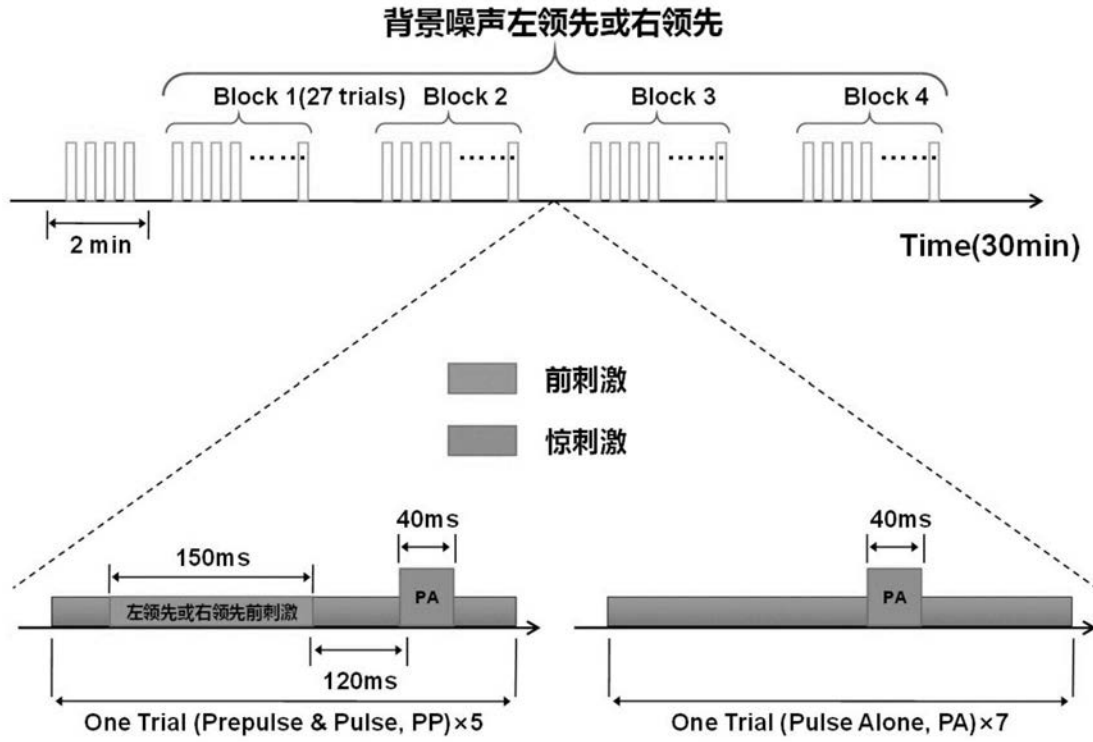


图2

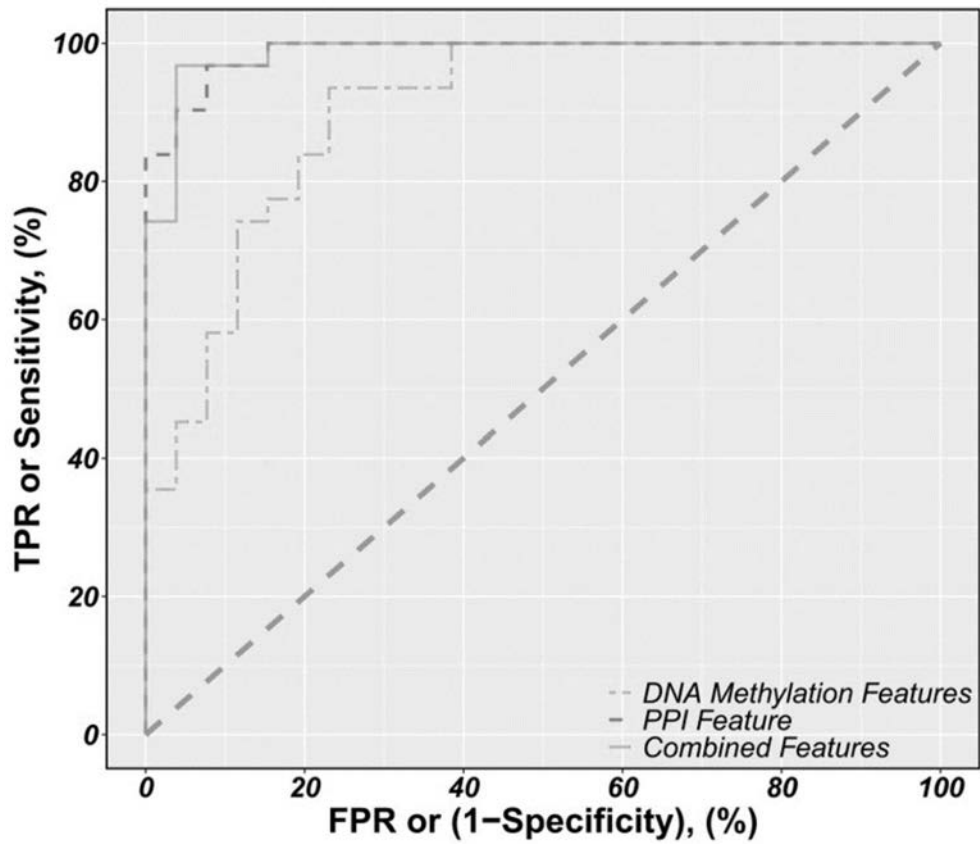


图3

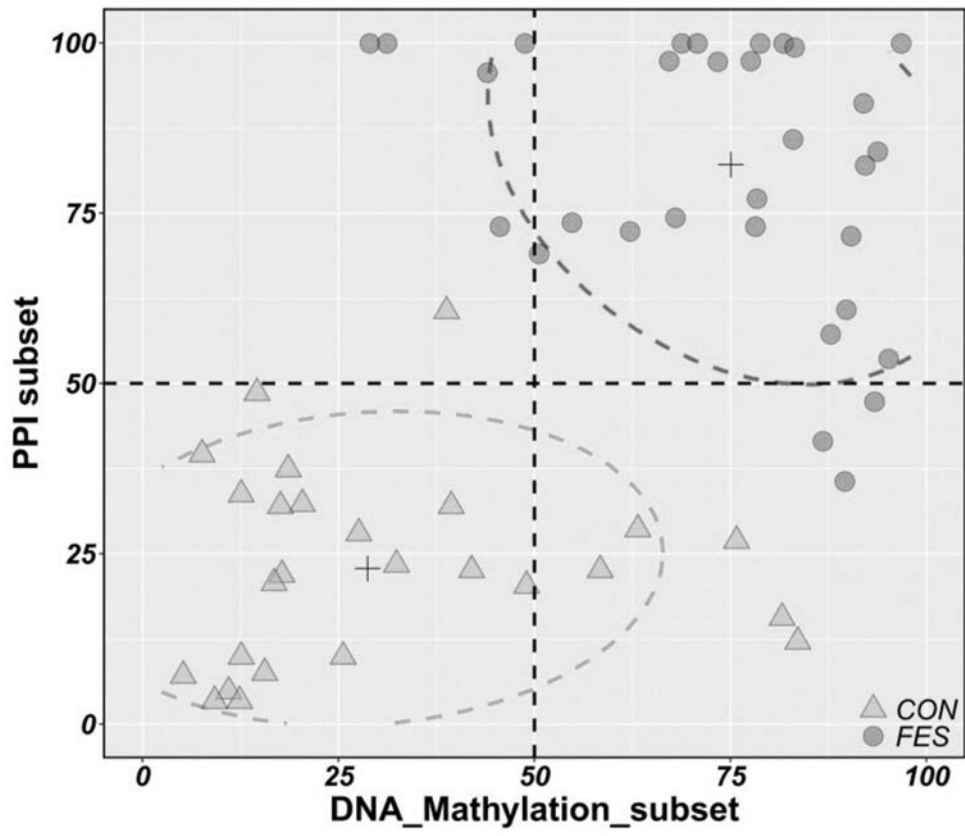


图4

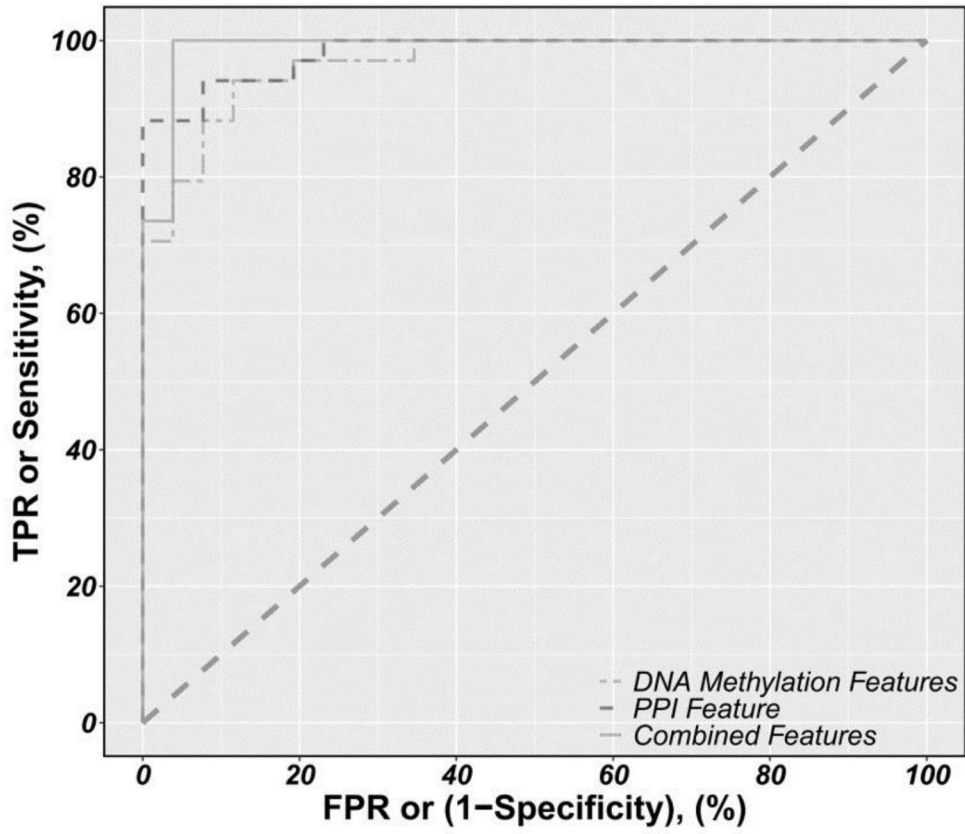


图5

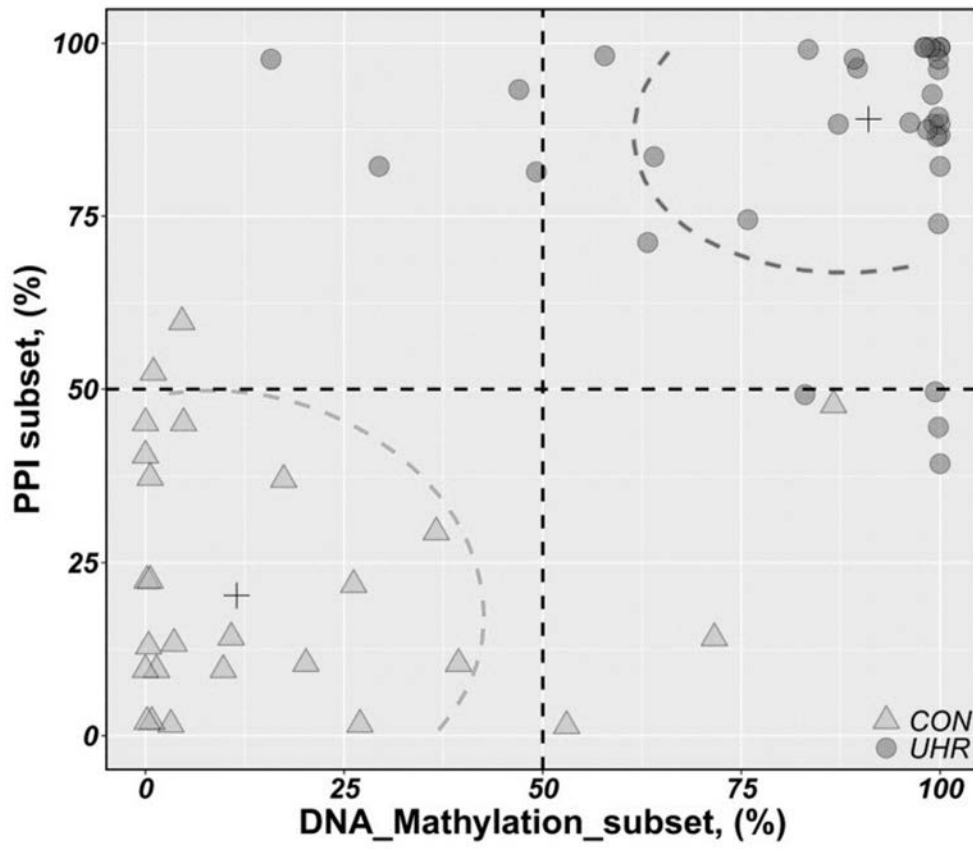


图6

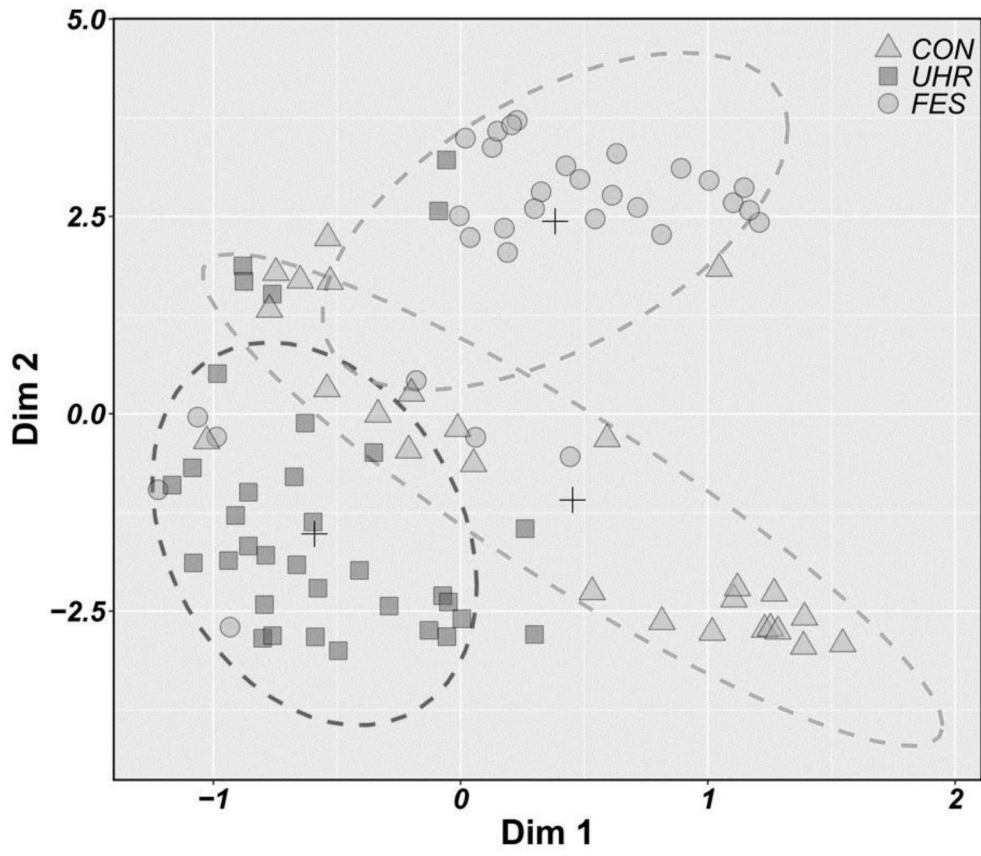


图7