(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0077997
(43) 공개일자 2019년07월04일

(51) 국제특허분류(Int. Cl.)
G16H 50/30 (2018.01) *A61B 5/00* (2006.01)
G16B 5/00 (2019.01)
(52) CPC특허분류
G16H 50/30 (2018.01)
A61B 5/7275 (2013.01)
(21) 출원번호 10-2017-0179635
(22) 출원일자 2017년12월26일
심사청구일자 2017년12월26일

(71) 출원인
주식회사 클리노믹스
울산광역시 울주군 언양읍 유니스트길 50, 경영관 806-6호
(72) 발명자
조윤성
경기도 수원시 영통구 광교로 114, 803호(이의동)
전재훈
경기도 수원시 팔달구 경수대로642번길 42-48, 201호(우만동)
(뒷면에 계속)
(74) 대리인
김순웅

전체 청구항 수 : 총 20 항

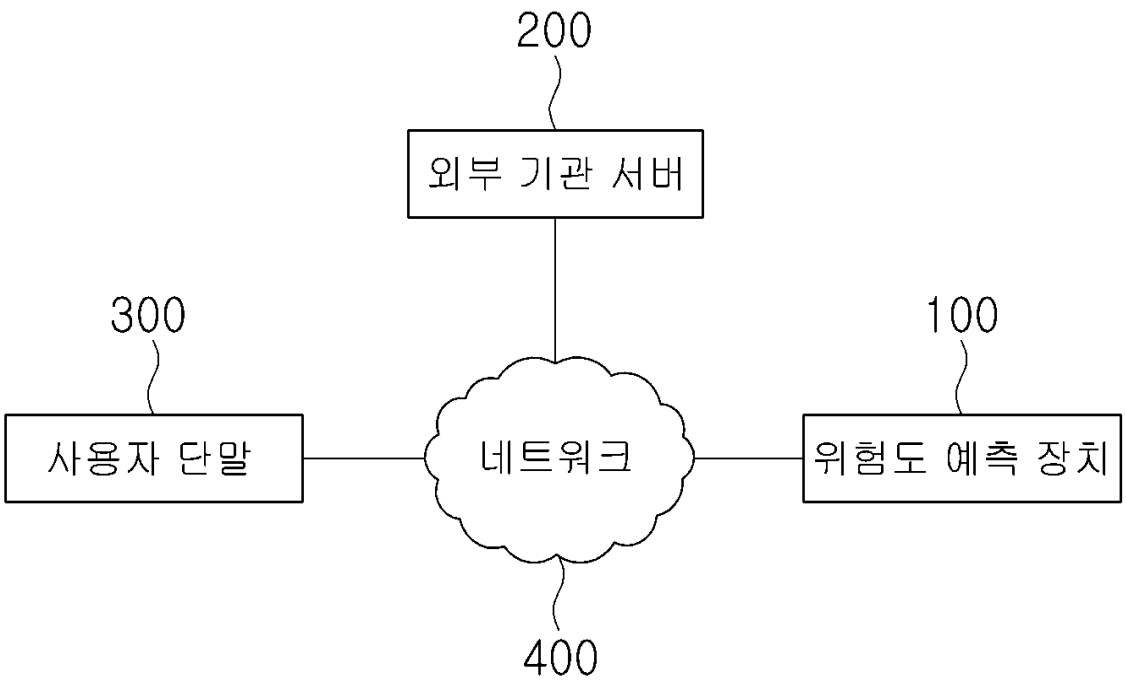
(54) 발명의 명칭 **질병 및 표현형 위험도 예측 장치 및 방법**

(57) 요 약

본 발명은 질병 및 표현형 위험도 예측 장치 및 방법에 관한 것으로서, 전 세계 인종/민족을 대상으로 개발된 여러 질병 및 표현형 관련 유전자 마커에 대해, 유전적 위험도를 검사하려는 집단에 적합하게 마커별 가중치를 부여하여 유전적 위험도를 예측함으로써 질병 및 표현형 위험도 예측의 정확도를 높일 수 있는 효과를 얻을 수

(뒷면에 계속)

대 표 도 - 도1



있다.

또한, 유전자검사를 통한 비질병 표현형 예측의 정확도를 높이기 위하여, 표현형에 미치는 증가영향과 감소영향을 구분하고 동시에 고려하여 유전적 위험도를 예측함으로써 질병 및 표현형 위험도 예측의 정확도를 높일 수 있는 효과를 얻을 수 있다.

또한, 의학적 장비를 통해 측정하지 않고, 질병 발생에 영향을 미치는 환경요인을 임상설문(Clinical Research)으로 수집하고, 수집된 환경적 요소를 유전적 요소와 통합하여 유전적 위험도를 예측함으로써 질병 및 표현형 위험도 예측의 정확도를 높일 수 있는 효과를 얻을 수 있다.

(52) CPC특허분류
G16B 99/00 (2019.02)

(72) 발명자
이황열
경기도 수원시 팔달구 세지로 279-1, 1016호(지동, 포레스트)
최진우
인천광역시 서구 여우재로75번길 18, 103동 103호 (가좌동, 한신빌라)
박종화
울산광역시 울주군 언양읍 유니스트길 50, 401동 601호(유니스트 교수아파트)
조수안
경기도 화성시 효행로291번길 26, 101동 1502호(기안동, 신일해피트리)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711038003
부처명 미래창조과학부
연구관리전문기관 정보통신산업진흥원
연구사업명 신산업창출을위한SW융합기술고도화
연구과제명 다중오믹스 기반 게놈 분석 가시화 플랫폼 개발
기 여 율 1/1
주관기관 (주)제로믹스
연구기간 2016.05.01 ~ 2017.12.31

명 세 서

청구범위

청구항 1

유전적 위험도 예측 장치로서,

외부 기관 서버로부터 질병에 대한 유전자 마커 정보를 획득하고, 사용자 단말로부터 사용자 유전형 정보를 획득하는 획득부;

유전자 마커 정보와 사용자 유전형 정보를 비교하여 질병에 대한 사용자의 위험인자 보유 수를 검출하고, 집단 내에서 위험인자의 빈도 값을 검출하고, 위험인자에 대한 유효크기와 사용자의 위험인자 보유 수를 이용하여 지노타입(genotype) 조합의 유효크기를 검출하는 검출부; 및

유전적 위험도를 산출하는 산출부;

를 포함하는, 유전적 위험도 예측 장치.

청구항 2

제1항에 있어서,

상기 산출부는,

수학식 $\sum w_i*F_i*R_i$을 이용하여 집단 내 대립유전자 빈도를 고려한 유전적 위험도를 산출하고,

이때, w는 위험인자에 대한 유효크기, F는 웨이트 값, R은 사용자의 위험인자 보유 수를 나타내고,

상기 웨이트 값은 상기 검출부에 의해 집단 내 위험인자 빈도 값이 변환된 값을 나타내며, 집단 내 위험인자 빈도 값이 0.5 이하인 경우, 웨이트 값은 집단 내 위험인자 빈도 값과 같고, 집단 내 위험인자 빈도 값이 0.5 초과인 경우, 웨이트 값은 1에서 집단 내 위험인자 빈도 값을 뺀 값인 것을 특징으로 하는, 유전적 위험도 예측 장치.

청구항 3

제1항에 있어서,

상기 산출부는,

수학식 $\prod W_i*F_i$을 이용하여 집단 내 대립유전자 빈도를 고려한 유전적 위험도를 산출하고,

이때, W는 지노타입 조합의 유효크기, F는 웨이트 값을 나타내고,

상기 웨이트 값은 상기 검출부에 의해 집단 내 위험인자 빈도 값이 변환된 값을 나타내며, 집단 내 위험인자 빈도 값이 0.5 이하인 경우, 웨이트 값은 집단 내 위험인자 빈도 값과 같고, 집단 내 위험인자 빈도 값이 0.5 초과인 경우, 웨이트 값은 1에서 집단 내 위험인자 빈도 값을 뺀 값인 것을 특징으로 하는, 유전적 위험도 예측 장치.

청구항 4

제1항에 있어서,

상기 검출부는,

상기 획득부로부터 획득된 유전자 마커 정보를 증가영향을 미치는 유전자 마커와 감소영향을 미치는 유전자 마커로 구별하고, 구별된 증가영향을 미치는 유전자 마커와 감소영향을 미치는 유전자 마커를 각각 사용자 유전형 정보와 비교하여 사용자의 증가영향 위험인자 보유 수와 감소영향 위험인자 보유 수를 검출하는, 유전적 위험도 예측 장치.

청구항 5

제4항에 있어서,

상기 산출부는,

수학식 $\sum W_{inc}*R_{inc} - \sum W_{dec}*R_{dec}$을 이용하여 표현형에 대한 증가 및 감소 영향을 고려한 유전적 위험도를 산출하며,

이때, W_{inc}는 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, W_{dec}는 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, R_{inc}는 사용자의 증가영향 위험인자 보유 수, R_{dec}는 사용자의 감소영향 위험인자 보유 수인, 유전적 위험도 예측 장치.

청구항 6

제4항에 있어서,

상기 산출부는,

수학식 $\prod W_{inc}*R_{inc} / \prod W_{dec}*R_{dec}$을 이용하여 표현형에 대한 증가 및 감소 영향을 고려한 유전적 위험도를 산출하며,

이때, W_{inc}는 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, W_{dec}는 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, R_{inc}는 사용자의 증가영향 위험인자 보유 수, R_{dec}는 사용자의 감소영향 위험인자 보유 수인, 유전적 위험도 예측 장치.

청구항 7

제6항에 있어서,

사용자의 증가영향 위험인자 보유 수가 0인 경우, 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기와 사용자의 증가영향 위험인자 보유 수를 곱한 값은 1이고,

사용자의 감소영향 위험인자 보유 수가 0인 경우, 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기와 사용자의 감소영향 위험인자 보유 수를 곱한 값은 1인 것을 특징으로 하는, 유전적 위험도 예측 장치.

청구항 8

제1항에 있어서,

상기 획득부는,

외부 기관 서버로부터 질병에 대한 유전적 요인과 환경적 요인의 비율을 획득하고, 사용자 단말로부터 사용자의 임상설문 정보를 획득하고,

상기 검출부는,

상기 획득부로부터 획득된 임상설문 정보를 이용하여 환경요인별 증가 위험도를 검출하는, 유전적 위험도 예측 장치.

청구항 9

제8항에 있어서,

상기 산출부는,

수학식 $GRS * (e_1*e_2*\cdots*e_n)$을 이용하여 유전적 요인과 환경적 요인을 고려한 통합 위험도를 산출하며,

이때, GRS는 유전적 요인을 이용하여 계산되는 유전적 위험도이고, e는 질병에 영향을 미치는 환경요인별 증가 위험도인, 유전적 위험도 예측 장치.

청구항 10

제8항에 있어서,

상기 산출부는,

수학식 $Pg*GRS + Pe*(e_1*e_2*\cdots*e_n)$을 이용하여 유전적 요인과 환경적 요인을 고려한 통합 위험도를 산출하며,

이때, GRS는 유전적 요인을 이용하여 계산되는 유전적 위험도이고, e는 질병에 영향을 미치는 환경요인별 증가 위험도이고, Pg는 유전적 요인이 질병에 영향을 미치는 비율이고, Pe는 환경적 요인이 질병에 영향을 미치는 비율인, 유전적 위험도 예측 장치.

청구항 11

유전적 위험도 예측 방법으로서,

외부 기관 서버로부터 질병에 대한 유전자 마커 정보를 획득하고, 사용자 단말로부터 사용자 유전형 정보를 획득하는 단계;

유전자 마커 정보와 사용자 유전형 정보를 비교하여 질병에 대한 사용자의 위험인자 보유 수를 검출하고, 집단 내에서 위험인자의 빈도 값을 검출하고, 위험인자에 대한 유효크기와 사용자의 위험인자 보유 수를 이용하여 지노타입(genotype) 조합의 유효크기를 검출하는 단계; 및

유전적 위험도를 산출하는 단계;

를 포함하는, 유전적 위험도 예측 방법.

청구항 12

제11항에 있어서,

상기 유전적 위험도는,

수학식 $\sum w_i*F_i*R_i$을 이용하여 계산되는 집단 내 대립유전자 빈도를 고려한 유전적 위험도이며,

이때, w는 위험인자에 대한 유효크기, F는 웨이트 값, R은 사용자의 위험인자 보유 수를 나타내고,

상기 웨이트 값은 상기 검출부에 의해 집단 내 위험인자 빈도 값이 변환된 값을 나타내며, 집단 내 위험인자 빈도 값이 0.5 이하인 경우, 웨이트 값은 집단 내 위험인자 빈도 값과 같고, 집단 내 위험인자 빈도 값이 0.5 초과인 경우, 웨이트 값은 1에서 집단 내 위험인자 빈도 값을 뺀 값인 것을 특징으로 하는, 유전적 위험도 예측 방법.

청구항 13

제11항에 있어서,

상기 유전적 위험도는,

수학식 $\prod W_i*F_i$을 이용하여 계산되는 집단 내 대립유전자 빈도를 고려한 유전적 위험도이며,

이때, W는 지노타입 조합의 유효크기, F는 웨이트 값을 나타내고,

상기 웨이트 값은 상기 검출부에 의해 집단 내 위험인자 빈도 값이 변환된 값을 나타내며, 집단 내 위험인자 빈도 값이 0.5 이하인 경우, 웨이트 값은 집단 내 위험인자 빈도 값과 같고, 집단 내 위험인자 빈도 값이 0.5 초과인 경우, 웨이트 값은 1에서 집단 내 위험인자 빈도 값을 뺀 값인 것을 특징으로 하는, 유전적 위험도 예측 방법.

청구항 14

제11항에 있어서,

상기 검출 단계는,

상기 획득 단계에서 획득된 유전자 마커 정보를 증가영향을 미치는 유전자 마커와 감소영향을 미치는 유전자 마커로 구별하고, 구별된 증가영향을 미치는 유전자 마커와 감소영향을 미치는 유전자 마커를 각각 사용자 유전형 정보와 비교하여 사용자의 증가영향 위험인자 보유 수와 감소영향 위험인자 보유 수를 검출하는 단계를 더 포함하는, 유전적 위험도 예측 방법.

청구항 15

제14항에 있어서,

상기 유전적 위험도는,

수학식 $\sum W_{inc}*R_{inc} - \sum W_{dec}*R_{dec}$을 이용하여 계산되는 표현형에 대한 증가 및 감소 영향을 고려한 유전적 위험도이며,

이때, W_{inc}는 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, W_{dec}는 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, R_{inc}는 사용자의 증가영향 위험인자 보유 수, R_{dec}는 사용자의 감소영향 위험인자 보유 수인, 유전적 위험도 예측 방법.

청구항 16

제14항에 있어서,

상기 유전적 위험도는,

수학식 $\prod W_{inc}*R_{inc} / \prod W_{dec}*R_{dec}$을 이용하여 계산되는 표현형에 대한 증가 및 감소 영향을 고려한 유전적 위험도이며,

이때, W_{inc}는 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, W_{dec}는 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, R_{inc}는 사용자의 증가영향 위험인자 보유 수, R_{dec}는 사용자의 감소영향 위험인자 보유 수인, 유전적 위험도 예측 방법.

청구항 17

제16항에 있어서,

사용자의 증가영향 위험인자 보유 수가 0인 경우, 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기와 사용자의 증가영향 위험인자 보유 수를 곱한 값은 1이고,

사용자의 감소영향 위험인자 보유 수가 0인 경우, 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기와 사용자의 감소영향 위험인자 보유 수를 곱한 값은 1인 것을 특징으로 하는, 유전적 위험도 예측 방법.

청구항 18

제11항에 있어서,

상기 획득 단계는,

외부 기관 서버로부터 질병에 대한 유전적 요인과 환경적 요인의 비율을 획득하고, 사용자 단말로부터 사용자의 임상설문 정보를 획득하는 단계를 더 포함하고,

상기 검출 단계는,

상기 획득 단계에서 획득된 임상설문 정보를 이용하여 환경요인별 증가 위험도를 검출하는 단계를 더 포함하는, 유전적 위험도 예측 방법.

청구항 19

제18항에 있어서,

상기 유전적 위험도는,

수학식 GRS * (e_1*e_2*…*e_n)을 이용하여 계산되는 유전적 요인과 환경적 요인을 고려한 통합 위험도이며,

이때, GRS는 유전적 요인을 이용하여 계산되는 유전적 위험도이고, e는 질병에 영향을 미치는 환경요인별 증가 위험도인, 유전적 위험도 예측 방법.

청구항 20

제18항에 있어서,

상기 유전적 위험도는,

수학식 Pg*GRS + Pe*(e_1*e_2*…*e_n)을 이용하여 계산되는 유전적 요인과 환경적 요인을 고려한 통합 위험도이며,

이때, GRS는 유전적 요인을 이용하여 계산되는 유전적 위험도이고, e는 질병에 영향을 미치는 환경요인별 증가 위험도이고, Pg는 유전적 요인이 질병에 영향을 미치는 비율이고, Pe는 환경적 요인이 질병에 영향을 미치는 비율인, 유전적 위험도 예측 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 질병 및 표현형 위험도 예측 장치 및 방법에 관한 것으로서, 기 보고된 질병 및 표현형 관련 유전자 마커들을 이용하여 질병 및 표현형 발생 가능성을 예측하는 데 있어, 1) 유전자 검사를 수행하려는 집단 내의 대립유전자 빈도, 2) 유전자 마커들이 표현형에 미치는 증가/감소 영향, 3) 환경요인을 고려하여 질병 및 표현형에 대한 위험도를 예측하는 질병 및 표현형 위험도 예측 장치 및 방법에 관한 것이다.

배 경 기 술

[0002] 유전체(Genome) 서열 해독 및 질병연구(Genome-wide association study; GWAS 등)에 의해, 다양한 질병 관련 혹은 표현형 관련 유전자 돌연변이(mutation) 마커가 밝혀지고 있다. 예를 들어, BRCA1, BRCA2 유전자의 돌연변이(mutation)의 유전체(Genome) 서열 해독 및 질병연구(Genome-wide association study; GWAS 등)에 의해, 다양한 질병 관련 혹은 표현형 관련 유전자 돌연변이(mutation) 마커가 밝혀지고 있다. 예를 들어, BRCA1, BRCA2 유전자의 돌연변이(mutation)의 경우 유방암 등의 질병 발생가능성을 증가시킨다고 보고되고 있다.

[0003] 그에 따라, 기존의 연구들을 통해 보고된 여러 유전자 마커를 활용하여 피검사자의 질병 및 표현형 발생 가능성을 예측하는 다양한 방법이 보고되고 있고, 국내외 유전자 검사기관에서도 유사한 방법을 활용하여 질병 및 표현형에 대한 유전적 위험도를 검사해주고 있다.

[0004] 현재 활용되고 있는 유전적 위험도 예측 방법에는, 특정 질병과 관련된 통계적으로 유의하다고 보고된 유전자 마커들을 대상으로, 테스트하는 시료의 유전자 위치를 해독 또는 지노타이핑하여, 위험 대립형질(Risk allele)을 몇 개 보유하고 있는지를 산술적으로 더하여 위험도를 계산하는 방법과 유전자 마커별 기 보고된 유효 크기(Effect size)를 위험인자에 가중치를 주어 산술합을 구하는 방법이 있다.

[0005] 그러나, 다양한 인종 및 집단별로 보유하고 있는 유전체(Genome) 서열은 집단 계층화(population stratification) 되어 있기 때문에, 질병 및 표현형에 영향을 미치는 유전자 마커 돌연변이 또한 인종 및 집단별로 돌연변이율이 상이하다. 따라서, 특정 유전자 마커는 특정 집단에서만 통계적으로 유의한 의미를 갖는 등 각 집단별 유효한 유전자 마커가 상이하다는 문제점이 있다.

[0006] 또한, 질병 관련 된 유전자 마커 개발의 경우 대부분 질병 발생을 높이는 것과 유의성을 갖는 유전자 마커들이 규명되어 있는데 반해, 비질병 표현형(지능, 키, 몸무게, 혈중 콜레스테롤 농도 등)의 경우 해당 표현형을 높이는 것과 낮추는 것이 구분되어 유전자 마커가 개발되어 있다. 예를 들어, Chan Y *et al*. Am. J. Hum. Genet. 2015 논문에서는, 앉은키(sitting height ratio)와 유의성을 갖는 유전자 마커를 개발하였는데, 큰 앉은키와 관련된 유전자 마커와 작은 앉은키와 관련된 유전자 마커가 동시에 구별되어 개발되었다. 이러한 증가/감소 영향을 고려하여 비질병 표현형을 예측하기 위해, 일반적으로 사용되는 방법은 한쪽 영향(증가 영향)에 대한 마커만을 활용하여, 한쪽 방향 기준으로 유전적 위험도를 예측해주는 것이다. 또 다른 방법으로는, 한쪽 방향(증가 방향)을 기준으로 유전적 위험도를 예측해주되, 반대 방향(감소 방향)으로 영향을 미치는 유전자 마커들은 역의

개념으로 감소 영향을 미치는 대립유전자(allele)가 아닌 다른 정상 대립유전자를 증가영향을 미치는 것으로 치환하여 유전적 위험도 예측에 유전자 마커로써 활용하는 방법이다. 그러나 이러한 방법은, 비질병 표현형에 대해 선택된 한 방향으로만 결과를 예측해준다는 문제점이 있으며, 한쪽 방향(증가)으로 영향을 미치는 대립유전자의 다른 정상 대립유전자가 반드시 반대 방향(감소)으로 영향을 미친다는 것을 의미하지는 않기 때문에, 표현형을 예측하는데 부정확함이 발생할 수 있다는 문제점이 있다.

[0007] 또한, 특정 유전적 영향에 의해 100% 발생하는 질병을 제외하고, 대부분의 질병은 유전적 요인 이외에 생활습관 등의 환경적 요인에 의해서도 질병 발생이 영향을 받게 된다. 예를 들어, 흡연 및 음주를 하거나, 방사능에 노출되는 경우, 보유하고 있는 유전적 특성과 함께 또는 별개로 특정 암에 대한 발생 가능성이 높아지게 된다. 최근, 유전적 요인과 비유전적 요인을 동시에 고려한 질병 예측/진단에 대한 연구가 발표되었는데, Goldstein BA *et al*. Front. Genet. 2014 논문에서는 관동맥성심장병(coronary heart disease)을 대상으로, 임상 위험 점수(clinical risk score)와 유전적 위험 점수(genetic risk score)를 로그-링크 함수(log-link function) 방법을 이용하여 질병발생을 예측하였다. 그러나 이 방법은 관동맥성심장병에 대해 임상적으로 위험도를 측정/판단하기 위한, Framingham Risk Score (성별, 연령, 콜레스테롤 수준, 혈압, 당뇨, 흡연상태 등을 고려)가 이미 정립이 되어 있었기에 적용 가능한 방법이었을 뿐더러, Framingham Risk Score 계산에 필요한 수치들은 건강검진/의료검사 등 의학적 장비로 측정한 정보가 필요하다는 문제점이 있다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 대한민국 등록특허공보 10-1582723(등록일자 2015년12월29일)

발명의 내용

해결하려는 과제

[0009] 본 발명은 상기 문제점을 해결하기 위한 것으로서, 전 세계 인종/민족을 대상으로 개발된 여러 질병 및 표현형 관련 유전자 마커에 대해, 유전적 위험도를 검사하려는 집단에 적합하게 마커별 가중치를 부여하여 유전적 위험도를 예측하는 질병 및 표현형 위험도 예측 장치 및 방법을 제공하는 데 그 목적이 있는 것이다.

[0010] 또한, 유전자검사를 통한 비질병 표현형 예측의 정확도를 높이기 위하여, 표현형에 미치는 증가영향과 감소영향을 구분하고 동시에 고려하여 유전적 위험도를 예측하는 질병 및 표현형 위험도 예측 장치 및 방법을 제공하는 데 그 목적이 있는 것이다.

[0011] 또한, 의학적 장비를 통해 측정하지 않고, 질병 발생에 영향을 미치는 환경요인을 임상설문(Clinical Research)을 통해 수집하고, 수집된 환경적 요소를 유전적 요소와 통합하여 유전적 위험도를 예측하는 질병 및 표현형 위험도 예측 장치 및 방법을 제공하는 데 그 목적이 있는 것이다.

과제의 해결 수단

[0012] 상기의 목적을 달성하기 위한 본 발명에 따른 질병 및 표현형 위험도 예측 장치는, 유전적 위험도 예측 장치로서, 외부 기관 서버로부터 질병에 대한 유전자 마커 정보를 획득하고, 사용자 단말로부터 사용자 유전형 정보를 획득하는 획득부; 유전자 마커 정보와 사용자 유전형 정보를 비교하여 질병에 대한 사용자의 위험인자 보유 수를 검출하고, 집단 내에서 위험인자의 빈도 값을 검출하고, 위험인자에 대한 유효크기와 사용자의 위험인자 보유 수를 이용하여 지노타입(genotype) 조합의 유효크기를 검출하는 검출부; 및 유전적 위험도를 산출하는 산출부;를 포함한다.

[0013] 산출부는, 수학식 $\sum w_i*F_i*R_i$을 이용하여 집단 내 대립유전자 빈도를 고려한 유전적 위험도를 산출할 수 있으며, 이때, w는 위험인자에 대한 유효크기, F는 웨이트 값, R은 사용자의 위험인자 보유 수를 나타낸다.

[0014] 또한, 산출부는, 수학식 $\prod W_i*F_i$을 이용하여 집단 내 대립유전자 빈도를 고려한 유전적 위험도를 산출할 수도 있으며, 이때, W는 지노타입 조합의 유효크기, F는 웨이트 값을 나타낸다.

[0015] 여기서, 상기 웨이트 값(F)은 상기 검출부에 의해 집단 내 위험인자 빈도 값이 변환된 값을 나타내며, 집단 내

위험인자 빈도 값이 0.5 이하인 경우, 웨이트 값은 집단 내 위험인자 빈도 값과 같고, 집단 내 위험인자 빈도 값이 0.5 초과인 경우, 웨이트 값은 1에서 집단 내 위험인자 빈도 값을 뺀 값인 것을 특징으로 한다.

[0016] 검출부는, 획득부로부터 획득된 유전자 마커 정보를 증가영향을 미치는 유전자 마커와 감소영향을 미치는 유전자 마커로 구별하고, 구별된 증가영향을 미치는 유전자 마커와 감소영향을 미치는 유전자 마커를 각각 사용자 유전형 정보와 비교하여 사용자의 증가영향 위험인자 보유 수와 감소영향 위험인자 보유 수를 검출한다.

[0017] 산출부는, 수학식 $\sum w_{inc}*R_{inc} - \sum w_{dec}*R_{dec}$을 이용하여 표현형에 대한 증가 및 감소 영향을 고려한 유전적 위험도를 산출할 수 있으며, 이때, w_{inc}는 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, w_{dec}는 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, R_{inc}는 사용자의 증가영향 위험인자 보유 수, R_{dec}는 사용자의 감소영향 위험인자 보유 수를 나타낸다.

[0018] 또한, 산출부는, 수학식 $\prod w_{inc}*R_{inc} / \prod w_{dec}*R_{dec}$을 이용하여 표현형에 대한 증가 및 감소 영향을 고려한 유전적 위험도를 산출할 수도 있으며, 이때, w_{inc}는 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, w_{dec}는 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, R_{inc}는 사용자의 증가영향 위험인자 보유 수, R_{dec}는 사용자의 감소영향 위험인자 보유 수를 나타낸다. 상기 수학식에서, 사용자의 증가영향 위험인자 보유 수가 0인 경우, 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기와 사용자의 증가영향 위험인자 보유 수를 곱한 값은 1이고, 사용자의 감소영향 위험인자 보유 수가 0인 경우, 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기와 사용자의 감소영향 위험인자 보유 수를 곱한 값은 1인 것을 특징으로 한다.

[0019] 획득부는, 외부 기관 서버로부터 질병에 대한 유전적 요인과 환경적 요인의 비율을 획득하고, 사용자 단말로부터 사용자의 임상설문 정보를 획득하고, 검출부는, 획득부로부터 획득된 임상설문 정보를 이용하여 환경요인별 증가 위험도를 검출한다.

[0020] 산출부는, 수학식 $GRS * (e_1*e_2*\cdots*e_n)$을 이용하여 유전적 요인과 환경적 요인을 고려한 통합 위험도를 산출할 수 있으며, 이때, GRS는 유전적 요인을 이용하여 계산되는 유전적 위험도이고, e는 질병에 영향을 미치는 환경요인별 증가 위험도를 나타낸다.

[0021] 또한, 산출부는, 수학식 $Pg*GRS + Pe*(e_1*e_2*\cdots*e_n)$을 이용하여 유전적 요인과 환경적 요인을 고려한 통합 위험도를 산출할 수도 있으며, 이때, GRS는 유전적 요인을 이용하여 계산되는 유전적 위험도이고, e는 질병에 영향을 미치는 환경요인별 증가 위험도이고, Pg는 유전적 요인이 질병에 영향을 미치는 비율이고, Pe는 환경적 요인이 질병에 영향을 미치는 비율을 나타낸다.

[0022] 상기의 목적을 달성하기 위한 본 발명에 따른 질병 및 표현형 위험도 예측 방법은, 유전적 위험도 예측 방법으로서, 외부 기관 서버로부터 질병에 대한 유전자 마커 정보를 획득하고, 사용자 단말로부터 사용자 유전형 정보를 획득하는 단계; 유전자 마커 정보와 사용자 유전형 정보를 비교하여 질병에 대한 사용자의 위험인자 보유 수를 검출하고, 집단 내에서 위험인자의 빈도 값을 검출하고, 위험인자에 대한 유효크기와 사용자의 위험인자 보유 수를 이용하여 지노타입(genotype) 조합의 유효크기를 검출하는 단계; 및 유전적 위험도를 산출하는 단계;를 포함한다.

[0023] 유전적 위험도는, 수학식 $\sum w_i*F_i*R_i$을 이용하여 계산되는 집단 내 대립유전자 빈도를 고려한 유전적 위험도일 수 있으며, 이때, w는 위험인자에 대한 유효크기, F는 웨이트 값, R은 사용자의 위험인자 보유 수를 나타낸다.

[0024] 또한, 유전적 위험도는, 수학식 $\prod W_i*F_i$을 이용하여 계산되는 집단 내 대립유전자 빈도를 고려한 유전적 위험도일 수도 있으며, 이때, W는 지노타입 조합의 유효크기, F는 웨이트 값을 나타낸다.

[0025] 여기서, 웨이트 값(F)은 검출부에 의해 집단 내 위험인자 빈도 값이 변환된 값을 나타내며, 집단 내 위험인자 빈도 값이 0.5 이하인 경우, 웨이트 값은 집단 내 위험인자 빈도 값과 같고, 집단 내 위험인자 빈도 값이 0.5 초과인 경우, 웨이트 값은 1에서 집단 내 위험인자 빈도 값을 뺀 값인 것을 특징으로 한다.

[0026] 검출 단계는, 획득 단계에서 획득된 유전자 마커 정보를 증가영향을 미치는 유전자 마커와 감소영향을 미치는 유전자 마커로 구별하고, 구별된 증가영향을 미치는 유전자 마커와 감소영향을 미치는 유전자 마커를 각각 사용자 유전형 정보와 비교하여 사용자의 증가영향 위험인자 보유 수와 감소영향 위험인자 보유 수를 검출하는 단계를 더 포함한다.

[0027] 유전적 위험도는, 수학식 $\sum w_{inc}*R_{inc} - \sum w_{dec}*R_{dec}$을 이용하여 계산되는 표현형에 대한 증가 및 감소 영향을 고려한 유전적 위험도일 수 있으며, 이때, w_{inc}는 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, w_{dec}는 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, R_{inc}는 사용자의 증가영향 위험인자 보유 수, R_{dec}는 사용자의 감소영향 위험인자 보유 수를 나타낸다.

[0028] 또한, 유전적 위험도는, 수학식 $\prod w_{inc}*R_{inc} / \prod w_{dec}*R_{dec}$을 이용하여 계산되는 표현형에 대한 증가 및 감소 영향을 고려한 유전적 위험도일 수도 있으며, 이때, w_{inc}는 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, w_{dec}는 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, R_{inc}는 사용자의 증가영향 위험인자 보유 수, R_{dec}는 사용자의 감소영향 위험인자 보유 수를 나타낸다. 상기 수학식에서, 사용자의 증가영향 위험인자 보유 수가 0인 경우, 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기와 사용자의 증가영향 위험인자 보유 수를 곱한 값은 1이고, 사용자의 감소영향 위험인자 보유 수가 0인 경우, 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기와 사용자의 감소영향 위험인자 보유 수를 곱한 값은 1인 것을 특징으로 한다.

[0029] 획득 단계는, 외부 기관 서버로부터 질병에 대한 유전적 요인과 환경적 요인의 비율을 획득하고, 사용자 단말로부터 사용자의 임상설문 정보를 획득하는 단계를 더 포함하고, 검출 단계는 획득 단계에서 획득된 임상설문 정보를 이용하여 환경요인별 증가 위험도를 검출하는 단계를 더 포함한다.

[0030] 유전적 위험도는, 수학식 $GRS * (e_1*e_2*\cdots*e_n)$을 이용하여 계산되는 유전적 요인과 환경적 요인을 고려한 통합 위험도일 수 있으며, 이때, GRS는 유전적 요인을 이용하여 계산되는 유전적 위험도이고, e는 질병에 영향을 미치는 환경요인별 증가 위험도를 나타낸다.

[0031] 또한, 유전적 위험도는, 수학식 $Pg*GRS + Pe*(e_1*e_2*\cdots*e_n)$을 이용하여 계산되는 유전적 요인과 환경적 요인을 고려한 통합 위험도일 수도 있으며, 이때, GRS는 유전적 요인을 이용하여 계산되는 유전적 위험도이고, e는 질병에 영향을 미치는 환경요인별 증가 위험도이고, Pg는 유전적 요인이 질병에 영향을 미치는 비율이고, Pe는 환경적 요인이 질병에 영향을 미치는 비율을 나타낸다.

발명의 효과

[0032] 본 발명에 따르면, 전 세계 인종 / 민족을 대상으로 개발된 여러 질병 및 표현형 관련 유전자 마커에 대해, 유전적 위험도를 검사하려는 집단에 적합하게 마커별 가중치를 부여하여 유전적 위험도를 예측함으로써 질병 및 표현형 위험도 예측의 정확도를 높일 수 있는 효과를 얻을 수 있다.

[0033] 또한, 유전자검사를 통한 비질병 표현형 예측의 정확도를 높이기 위하여, 표현형에 미치는 증가영향과 감소영향을 구분하고 동시에 고려하여 유전적 위험도를 예측함으로써 질병 및 표현형 위험도 예측의 정확도를 높일 수 있는 효과를 얻을 수 있다.

[0034] 또한, 의학적 장비를 통해 측정하지 않고, 질병 발생에 영향을 미치는 환경요인을 임상설문(Clinical Research)으로 수집하고, 수집된 환경적 요소를 유전적 요소와 통합하여 유전적 위험도를 예측함으로써 질병 및 표현형 위험도 예측의 정확도를 높일 수 있는 효과를 얻을 수 있다.

도면의 간단한 설명

[0035] 도 1은 본 발명에 따른 유전적 위험도 예측 장치가 포함된 시스템 구성도이다.

도 2는 도 1에 도시된 유전적 위험도 예측 장치의 구성을 보다 상세히 나타낸 블록도이다.

도 3은 본 발명의 제 1 실시예에 따른 질병 및 표현형 위험도 예측 방법을 설명하기 위한 흐름도이다.

도 4는 본 발명의 제 2 실시예에 따른 질병 및 표현형 위험도 예측 방법을 설명하기 위한 흐름도이다.

도 5는 본 발명의 제 3 실시예에 따른 질병 및 표현형 위험도 예측 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0036] 이하에서 첨부한 도면을 참조하여 본 발명에 따른 질병 및 표현형 위험도 예측 장치 및 방법의 바람직한 실시예에 대해 상세하게 설명한다.

[0037] 먼저, 도 1 및 도 2를 참조하여, 본 발명에 따른 유전적 위험도 예측 장치에 대하여 설명한다.

[0038] 도 1은 본 발명에 따른 유전적 위험도 예측 장치가 포함된 시스템 구성도이다.

[0039] 도 1에 도시된 바와 같이, 본 발명에 따른 유전적 위험도 예측 장치(이하 '위험도 예측 장치'라 한다)(100)는 네트워크(400)를 통해 외부 기관 서버(200) 및 사용자 단말(300)과 연결된다.

[0040] 위험도 예측 장치(100)는 외부 기관 서버(200)로부터 제공받은 각종 데이터, 예를 들면 기 구축된 질병 관련 통계 정보 및 기 구축된 유전자 마커 정보 등과, 사용자 단말(300)로부터 제공받은 각종 데이터, 예를 들면 사용자의 유전 정보 등을 이용하여, 질병에 대한 유전적 위험도를 계산한다.

[0042] 외부 기관 서버(200)는 병원 등의 의료 서비스 제공 기관, 보건복지부 등의 의료 국가기관, 건강보험심사평가원 등의 의료 공공기관 등의 외부 기관에서 운영하는 서버를 말한다. 외부 기관 서버(200)는 사용자의 진료 정보, 기 구축된 질병 관련 통계 정보, 기 구축된 질병에 대한 유전자 마커 정보 등을 네트워크(400)를 통해 위험도 예측 장치(100)로 제공할 수 있다.

[0043] 여기서, 사용자의 진료 정보는 사용자가 의료 서비스 제공 기관에서 진료받은 이력을 나타내는 정보로서, 사용자의 이름, 사용자 식별 정보(예를 들면, 주민등록번호 등), 사용자가 진료받은 의료 서비스 제공 기관의 식별 정보, 진료받은 의사, 진료받은 날짜, 진료 내용(예를 들면, 사용자의 보유 질병 등) 등을 포함할 수 있다.

[0044] 그리고, 질병 관련 통계 정보는 의료 국가기관이나 의료 공공기관에서 보건의료 정책 수립 등의 이유로 구축된 질병과 관련된 다양한 통계 정보를 말한다. 예를 들면, 질병 관련 통계 정보는 연령에 따른 질병 관련 통계 자료, 성별에 따른 질병 관련 통계 자료, 직업에 따른 질병 관련 통계 자료, 지역에 따른 질병 통계 자료 등을 포함할 수 있다.

[0045] 또한, 유전자 마커 정보는 유전자 변이를 일으키는 질병 관련 인자를 갖는 유전자 정보로서, 사용자의 질병 관련 위험도를 예측하기 위해 사용자 유전 정보와 대조되는 정보이다. 유전자 마커 정보는 유전자 변이 위치, 대립인자, 질병 위험인자, 위험인자에 대한 유효크키 등을 포함할 수 있다.

[0047] 사용자 단말(300)은 네트워크(400)를 통해 위험도 예측 장치(100)와 연결되어 각종 데이터를 주고받을 수 있다.

[0048] 즉, 사용자 단말(300)은 사용자의 개인 정보, 사용자의 유전 정보 등을 네트워크(400)를 통해 위험도 예측 장치(100)로 제공할 수 있다. 여기서, 사용자의 개인 정보는 사용자의 이름, 사용자의 식별 정보(예를 들면, 주민등록번호 등), 사용자의 성별, 사용자의 연령, 사용자의 거주지, 사용자의 직업, 사용자의 직장 주소 등을 포함할 수 있다. 그리고, 사용자의 유전 정보(genetic information)는 유전자 변이를 일으키는 질병 위험인자와 대조하기 위한 정보로서, DNA 염기서열을 이루는 아데닌(adenine, A), 티민(thymine, T), 시토신(cytosine, C), 구아닌(guanine, G) 중 두 개로 구성된 유전 정보이다. 예를 들면, AT, CG, AA, CC, TT, GG 등이다.

[0049] 그리고, 사용자 단말(300)은 네트워크(400)를 통해 위험도 예측 장치(100)로부터 제공받은 정보를 디스플레이할 수 있다.

[0050] 여기서, 사용자 단말(300)은 데스크톱 컴퓨터뿐만 아니라 노트북 컴퓨터, 워크스테이션, 팜톱(palmtop) 컴퓨터, UMPC(Ultra Mobile Personal Computer), 태블릿 PC, 개인 휴대 정보 단말기(Personal Digital Assistant, PDA), 웹 패드, 스마트폰, 휴대전화 등과 같이 메모리 수단을 구비하고 마이크로 프로세서를 탑재하여 연산 능력을 갖춘 단말기로 이루어질 수 있다.

[0052] 네트워크(400)는 구내 정보 통신망(LAN), 도시권 통신망(MAN), 광역 통신망(WAN) 및 인터넷 등을 포함하는 데이터 통신망뿐만 아니라 방송망, 전화망 등을 포함할 수 있고, 유선과 무선을 가리지 않으며, 모든 통신 방식이 사용 가능하다.

[0054] 도 2는 도 1에 도시된 유전적 위험도 예측 장치의 구성을 보다 상세히 나타낸 블록도이다.

[0055] 도 2에 도시된 바와 같이, 위험도 예측 장치(100)는 획득부(120), 검출부(140) 및 산출부(160)를 포함할 수 있

다. 위험도 예측 장치(100)는 세 가지 실시예에 따라 유전적 위험도를 예측할 수 있으며, 상기 세 가지 실시예들은 동시 또는 선택적으로 통합하여 활용 가능하다.

[0056] 본 발명의 제 1 실시예에 따르면, 위험도 예측 장치(100)는 전 세계 인종/민족을 대상으로 기 보고된 질병 및 표현형 관련 유전자 마커에 대해, 유전적 위험도를 검사하려는 집단내 대립유전자 빈도(allele frequency) 정보를 고려하여 해당 집단에 적합하게 마커별 가중치를 부여하고 유전적 위험도를 예측한다.

[0057] 본 발명의 제 1 실시예에 따르면, 획득부(120)는 외부 기관 서버(200)로부터 질병에 대한 유전자 마커 정보를 획득하고, 사용자 단말(300)로부터 사용자 유전형 정보를 획득한다.

[0058] 이때, 유전자 마커 정보는 유전자 변이 위치, 대립인자, 질병 위험인자, 위험인자에 대한 오즈비(odds ratio)로 나타낸 유효크기 등을 포함하는 정보이며, 사용자의 유전형 정보는 유전자 변이를 일으키는 질병 위험인자와 대조하기 위한 정보로서, DNA 염기서열을 이루는 아데닌(adenine, A), 티민(thymine, T), 시토신(cytosine, C), 구아닌(guanine, G) 중 두 개로 구성된 유전 정보이다.

[0059] 본 발명의 제 1 실시예에 따르면, 검출부(140)는 유전자 마커 정보와 사용자 유전형 정보를 비교하여 질병에 대한 사용자의 위험인자 보유 수를 검출하고, 집단 내에서 위험인자의 빈도 값을 검출하고, 위험인자에 대한 유효크기와 사용자의 위험인자 보유 수를 이용하여 지노타입(genotype) 조합의 유효크기를 검출한다.

[0060] 여기서, 집단 내 위험인자 빈도 값은 사용자가 속한 집단 내에서의 해당 위험인자의 보유 빈도를 말하며, 각각의 유전자 마커에 가중치로서 활용된다. 이때, 집단 내 위험인자 빈도 값이 0.5에 가까울수록 높은 가중치를 갖게 된다.

[0061] 그리고, 지노타입 조합의 유효크기는 위험인자의 유효크기와 피검자의 해당 위험인자 보유 수에 따라 결정되는 값을 말한다.

[0062] 또한, 검출부(140)는 검출된 집단 내 위험인자 빈도 값을 웨이트(weight) 값으로 변환하는데, 이때, 집단 내 위험인자 빈도 값이 0.5 이하인 경우, 웨이트 값으로 변환한 값은 집단 내 위험인자 빈도 값과 같고, 집단 내 위험인자 빈도 값이 0.5 초과인 경우, 웨이트 값으로 변환한 값은 1에서 집단 내 위험인자 빈도 값을 뺀 값이다.

[0063] 본 발명의 제 1 실시예에 따르면, 산출부(160)는 사용자의 질병에 대한 유전적 위험도를 산출한다.

[0064] 이때, 유전적 위험도는 수학식 $\sum w_i*F_i*R_i$ 또는 $\sum \log(w_i)*R_i$을 이용하여 산출될 수 있으며, 여기서, w는 위험인자에 대한 유효크기이고, F는 웨이트 값으로 변환된 집단 내 위험인자의 빈도 값이며, R은 사용자의 위험인자 보유 수를 말한다.

[0065] 또한, 유전적 위험도는 수학식 $\prod W_i*F_i$을 이용하여 산출될 수도 있으며, 여기서, W는 지노타입 조합의 유효크기이고, F는 웨이트 값으로 변환된 집단 내 위험인자의 빈도 값을 말한다.

[0067] 본 발명의 제 2 실시예에 따르면, 위험도 예측 장치(100)는 유전자 검사를 통한 비질병 표현형 예측의 정확도를 높이기 위하여, 표현형에 미치는 증가영향과 감소영향을 구분하고 동시에 고려하여 유전적 위험도를 예측한다.

[0068] 본 발명의 제 2 실시예에 따르면, 획득부(120)는 외부 기관 서버(200)로부터 질병에 대한 유전자 마커 정보를 획득하고, 사용자 단말(300)로부터 사용자 유전형 정보를 획득한다.

[0069] 이때, 유전자 마커 정보는 유전자의 변이 위치, 질병 위험인자, 위험인자에 대한 베타 계수(beta coefficient)로 나타낸 유효크기 등을 포함하는 정보이며, 사용자의 유전형 정보는 유전자 변이를 일으키는 질병 위험인자와 대조하기 위한 정보로서, DNA 염기서열을 이루는 아데닌(adenine, A), 티민(thymine, T), 시토신(cytosine, C), 구아닌(guanine, G) 중 두 개로 구성된 유전 정보이다.

[0070] 본 발명의 제 2 실시예에 따르면, 검출부(140)는 특정 표현형과 관련된 통계적으로 유의한 유전자 마커 정보를 증가영향을 미치는 유전자 마커와 감소영향을 미치는 유전자 마커로 구별하고, 구별된 증가영향을 미치는 유전자 마커와 감소영향을 미치는 유전자 마커를 각각 사용자 유전형 정보와 비교하여 사용자가 증가영향 위험인자와 감소영향 위험인자를 각각 몇 개 보유하고 있는지를 검출한다.

[0071] 본 발명의 제 2 실시예에 따르면, 산출부(160)는 표현형에 대한 유전적 위험도를 산출한다.

[0072] 이때, 표현형에 대한 유전적 위험도는 수학식 $\sum w_{inc}*R_{inc} - \sum w_{dec}*R_{dec}$을 이용하여 산출될 수 있으며, 여기서, w_{inc}

는 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, w_{dec}는 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기, R_{inc}는 사용자의 증가영향 위험인자 보유 수, R_{dec}는 사용자의 감소영향 위험인자 보유 수를 말한다.

[0073] 여기서, 사용자의 증가영향 위험인자 보유 수와 감소영향 위험인자 보유 수는 0, 1, 2 중 하나의 값이다.

[0074] 또한, 표현형에 대한 유전적 위험도는 수학식 $\prod w_{inc}*R_{inc} / \prod w_{dec}*R_{dec}$을 이용하여 산출될 수도 있으며, 여기서, w_{inc}는 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기이고, w_{dec}는 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기이고, R_{inc}는 사용자의 증가영향 위험인자 보유 수이고, R_{dec}는 사용자의 감소영향 위험인자 보유 수를 말한다.

[0075] 여기서, 사용자의 증가영향 위험인자 보유 수와 감소영향 위험인자 보유 수는 1, 2 중 하나의 값이다. 만약 사용자의 증가영향 위험인자 보유 수와 감소영향 위험인자 보유 수가 0인 경우에는 상기 수학식에서 유효크기와 위험인자 보유 수를 곱한 값을 1로 하여 계산한다.

[0076] 즉, 사용자의 증가영향 위험인자 보유 수가 0인 경우, 증가영향을 미치는 유전자 마커의 위험인자에 대한 유효크기와 사용자의 증가영향 위험인자 보유 수를 곱한 값은 1이고, 사용자의 감소영향 위험인자 보유 수가 0인 경우, 감소영향을 미치는 유전자 마커의 위험인자에 대한 유효크기와 사용자의 감소영향 위험인자 보유 수를 곱한 값은 1이다.

[0078] 본 발명의 제 3 실시예에 따르면, 위험도 예측 장치(100)는 유전적 요인뿐만 아니라, 환경적 요인을 동시에 고려하여 유전적 위험도를 예측한다.

[0079] 본 발명의 제 3 실시예에 따르면, 획득부(120)는 외부 기관 서버(200)로부터 질병에 대한 유전자 마커 정보 및 유전적 요인과 환경적 요인의 질병 발생 비율을 획득하고, 사용자 단말(300)로부터 사용자 유전형 정보와 임상설문 정보를 획득한다.

[0080] 유전자 마커 정보는 유전자 변이를 일으키는 질병 관련 인자를 갖는 유전자 정보로서, 사용자의 질병 관련 위험도를 예측하기 위해 사용자 유전 정보와 대조되는 정보이다. 유전자 마커 정보는 유전자 변이 위치, 대립인자, 질병 위험인자, 위험인자에 대한 유효크키 등을 포함할 수 있다.

[0081] 사용자의 유전형 정보는 유전자 변이를 일으키는 질병 위험인자와 대조하기 위한 정보로서, DNA 염기서열을 이루는 아데닌(adenine, A), 티민(thymine, T), 시토신(cytosine, C), 구아닌(guanine, G) 중 두 개로 구성된 유전 정보이다.

[0082] 임상설문 정보는 사용자의 생활 습관에 관련된 설문 정보로서, 임상설문 정보를 통해 질병 발생에 영향을 미치는 환경 요인을 알 수 있다.

[0083] 본 발명의 제 3 실시예에 따르면, 검출부(140)는 유전자 마커 정보와 사용자 유전형 정보를 이용하여 유전적 위험도를 계산하고, 임상설문 정보를 이용하여 환경 요인별 증가 위험도를 검출한다.

[0084] 환경 요인별 증가 위험도는, 환경 요인(예를 들면, 흡연, 음주 등)에 따라 질병이 발생할 위험도를 나타낸다.

[0085] 본 발명의 제 3 실시예에 따르면, 산출부(160)는 유전적 / 환경적 요인 통합 위험도를 산출한다.

[0086] 이때, 유전적 / 환경적 요인 통합 위험도는 수학식 $GRS * (e_1*e_2*\cdots*e_n)$을 이용하여 산출될 수 있으며, 여기서, GRS는 유전적 위험도이고, e는 질병에 영향을 미치는 환경요인의 증가 위험도를 말한다.

[0087] GRS는 유전적 요인(예를 들면, 위험인자 등)을 이용하여 다양한 방법으로 계산되는 값으로서, 본 발명에 따른 제 3 실시예에서는 유전적 위험도를 계산하는 방법을 한정하는 것은 아니다.

[0088] 또한, 유전적 / 환경적 요인 통합 위험도는 수학식 $Pg*GRS + Pe*(e_1*e_2*\cdots*e_n)$을 이용하여 산출될 수도 있으며, 여기서, GRS는 유전적 위험도이고, e는 질병에 영향을 미치는 환경요인의 증가 위험도이고, Pg는 유전적 요인이 질병에 영향을 미치는 비율이고, Pe는 환경적 요인이 질병에 영향을 미치는 비율을 말한다.

[0090] 그러면, 도 3 내지 도 5를 참조하여 본 발명에 따른 질병 및 표현형 위험도 예측 방법을 상세하게 설명한다. 본 발명에 따른 위험도 예측 방법은 세 가지 실시예에 따라 유전적 위험도를 예측할 수 있으며, 상기 세 가지 실시예들은 동시 또는 선택적으로 통합하여 활용 가능하다.

[0091] 도 3은 본 발명의 제 1 실시예에 따른 질병 및 표현형 위험도 예측 방법을 설명하기 위한 흐름도이다. 본 발명의 제 1 실시예에 따르면, 전 세계 인종/민족을 대상으로 기 보고된 질병 및 표현형 관련 유전자 마커에 대해, 유전적 위험도를 검사하려는 집단내 대립유전자 빈도(allele frequency) 정보를 고려하여 해당 집단에 적합하게 마커별 가중치를 부여함으로써 유전적 위험도가 예측된다.

[0092] 본 발명의 제 1 실시예에 따르면, 외부 기관 서버(200)로부터 질병에 대한 유전자 마커 정보를 획득하고, 사용자 단말(300)로부터 사용자 유전형 정보를 획득한다(S10).

[0093] 이후, 유전자 마커 정보와 사용자 유전형 정보를 비교하여 질병에 대한 사용자의 위험인자 보유 수를 검출하고, 집단 내에서 위험인자의 빈도 값을 검출하고, 위험인자에 대한 유효크기와 사용자의 위험인자 보유 수를 이용하여 지노타입(genotype) 조합의 유효크기를 검출한다(S12).

[0094] 아래 [표 1]을 참고하면, 유전자의 위치 4q23의 변이 위험인자는 C인 것을 알 수 있다. 이를 피검자 1, 2, 3의 유전형과 비교해보면, 피검자 1의 위험인자 보유 수는 0개, 피검자 2의 위험인자 보유 수는 1개, 피검자 3의 위험인자 보유 수는 0개인 것을 알 수 있다.

[0095] 그리고, 검출된 위험인자 보유 수를 통해 피검자들의 지노타입 조합의 유효크기를 알 수 있다. [표 1]을 참고하면, 유전자의 위치 4q23의 위험인자 C에 대한 유효크기는 1.66이고, 피검자들의 위험인자 보유 수는 각각 0개, 1개 0개이다. 이를 이용하여 피검자들의 지노타입 조합의 유효크기를 계산하면, 피검자 1의 지노타입 조합의 유효크기는 1, 피검자 2의 지노타입 조합의 유효크기는 1.66, 피검자 3의 지노타입 조합의 유효크기는 1이다.

[0096] 이때, 위험인자 보유 수가 0인 경우, 지노타입 조합의 유효크기는 1인 것으로 한다.

[0097] 이후, 검출된 집단 내 위험인자 빈도 값을 웨이트(weight) 값으로 변환한다(S14). 이때, 집단 내 위험인자 빈도 값이 0.5 이하인 경우, 웨이트 값으로 변환된 값은 집단 내 위험인자 빈도 값과 같고, 집단 내 위험인자 빈도 값이 0.5 초과인 경우, 웨이트 값으로 변환된 값은 1에서 집단 내 위험인자 빈도 값을 뺀 값과 같다. [표 1]을 참고하면, 유전자의 위치 4q23의 변이 위험인자 C의 집단 내 보유 빈도는 0.21로서 0.5를 넘지 않으므로, 웨이트로 변환한 값은 0.21이 된다.

[0098] 이후, 수학식 $\sum w_i*F_i*R_i$ 또는 $\sum \log(w_i)*R_i$ 또는 수학식 $\prod W_i*F_i$을 이용하여 유전적 위험도를 산출한다(S16).

[0099] 예를 들면, [표 1]에 나타난 값들을 수학식 $\sum w_i*F_i*R_i$에 적용하여 유전적 위험도를 계산해보면, 피검자 1의 유전적 위험도는 (1.66*0.21*0)+(1.47*0.24*0)+(1.85*0.12*0)=0이고, 피검자 2의 유전적 위험도는 (1.66*0.21*1)+(1.47*0.24*0)+(1.85*0.12*1)=0.5706이고, 피검자 3의 유전적 위험도는 (1.66*0.21*0)+(1.47*0.24*2)+(1.85*0.12*0)=0.7056이다.

[0100] 또한, [표 1]에 나타난 값들을 수학식 $\prod W_i*F_i$에 적용하여 유전적 위험도를 계산해보면, 피검자 1의 유전적 위험도는 (1*0.21)*(1*0.24)*(1*0.12)=0.006048이고, 피검자 2의 유전적 위험도는 (1.66*0.21)*(1*0.24)*(1.85*0.12)=0.01857이고, 피검자 3의 유전적 위험도는 (1*0.21)*(2.16*0.24)*(1*0.12)=0.01306이다.

표 1

[0102]

유전자 마커 정보					피검자의 보유 유전형			지노타입 조합별 유효 크기		
유전자 변이 위치	대립 인자	위험 인자	유효 크기	집단 내 위험인자 빈도	피검자 1	피검자 2	피검자 3	피검자 1	피검자 2	피검자 3
4q23	C/T	C	1.66	0.21	TT	CT	TT	1	1.66	1
10q23	G/A	G	1.47	0.24	AA	AA	GG	1	1	2.16
12q24	A/G	A	1.85	0.12	GG	AG	GG	1	1.85	1

[0104] 도 4는 본 발명의 제 2 실시예에 따른 질병 및 표현형 위험도 예측 방법을 설명하기 위한 흐름도이다. 본 발명의 제 2 실시예에 따르면, 유전자 검사를 통한 비질병 표현형 예측의 정확도를 높이기 위하여, 표현형에 미치는 증가영향과 감소영향을 구분하고 동시에 고려함으로써 유전적 위험도가 예측된다.

[0105] 본 발명의 제 2 실시예에 따르면, 외부 기관 서버(200)로부터 질병에 대한 유전자 마커 정보를 획득하고, 사용자 단말(300)로부터 사용자 유전형 정보를 획득한다(S20).

[0106] 이후, 유전자 마커 정보를 증가영향을 미치는 유전자 마커와 감소영향을 미치는 유전자 마커로 구별하고(S22), 구별된 증가영향을 미치는 유전자 마커와 감소영향을 미치는 유전자 마커를 각각 사용자 유전형 정보와 비교하여(S24) 사용자의 증가영향 위험인자 보유 수와 감소영향 위험인자 보유 수를 검출한다(S26).

[0107] 아래 [표 2]를 참고하면, 증가영향을 미치는 유전자 마커는 rs10495928, rs1800562, rs10224002, rs16926246, rs4671393, rs9399137이고, 감소영향을 미치는 유전자 마커는 rs6013509, rs570013781, rs5030868이다. 예를 들면, 증가영향을 미치는 유전자 마커 rs10495928의 질병 위험인자는 G이고, 이를 피검자의 유전형 GT와 비교하면, 피검자는 증가영향을 미치는 위험인자를 1개 보유하고 있는 것을 알 수 있다.

[0108] 이후, 수학식 $\sum W_{inc}*R_{inc} - \sum W_{dec}*R_{dec}$ 또는 수학식 $\prod W_{inc}*R_{inc} / \prod W_{dec}*R_{dec}$을 이용하여 표현형에 대한 유전적 위험도를 산출한다(S28).

[0109] 예를 들면, [표 2]에 나타난 값들을 수학식 $\sum W_{inc}*R_{inc} - \sum W_{dec}*R_{dec}$에 적용하여 표현형에 대한 유전적 위험도를 계산해보면,

[0110] (0.06*1+0.16*2+0.07*1+0.11*1+0.58*0+0.42*2)-(0.06*1+0.20*2+0.13*1)=0.81이다.

[0111] 또한, [표 2]에 나타난 값들을 수학식 $\prod W_{inc}*R_{inc} / \prod W_{dec}*R_{dec}$에 적용하여 표현형에 대한 유전적 위험도를 계산해보면, [(0.06*1)*(0.16*2)*(0.07*1)*(0.11*1)*1*(0.42*2)]/[(0.06*1)*(0.20*2)*(0.13*1)]=0.040이다. 여기서, 사용자가 보유한 위험인자의 개수가 0인 경우, 즉, 사용자가 증가영향을 미치는 위험인자를 보유하고 있지 않거나, 감소영향을 미치는 위험인자를 보유하고 있지 않을 때, 값을 1로 적용하여 계산한다.

표 2

[0113]

유전자 변이위치	기보고된 위험인자	유효크기	증가 / 감소 영향	피검자의 유전형	증가영향 보유 위험인자 개수	감소영향 보유 위험인자 개수
rs10495928	G	0.06 g//dl	increase	GT	1	-
rs1800562	A	0.16 g//dl	increase	AA	2	-
rs10224002	G	0.07 g//dl	increase	GA	1	-
rs16926246	T	0.11 g//dl	increase	TA	1	-
rs4671393	A	0.58 g//dl	increase	TT	0	-
rs9399137	C	0.42 g//dl	increase	CC	2	-
rs6013509	A	0.06 g//dl	decrease	AT	-	1
rs570013781	A	0.20 g//dl	decrease	AA	-	2
rs5030868	A	0.13 g//dl	decrease	AG	-	1
합계	-		-		7	4

[0115] 도 5는 본 발명의 제 3 실시예에 따른 질병 및 표현형 위험도 예측 방법을 설명하기 위한 흐름도이다. 본 발명의 제 3 실시예에 따르면, 유전적 요인뿐만 아니라, 환경적 요인을 동시에 고려함으로써 유전적 위험도가 예측된다.

[0116] 본 발명의 제 3 실시예에 따르면, 외부 기관 서버(200)로부터 질병에 대한 유전자 마커 정보 및 질병에 대한 유전적 요인과 환경적 요인의 비율을 획득하고, 사용자 단말(300)로부터 사용자 유전형 정보와 임상설문 정보를

획득한다(S30).

[0117] 이후, 유전자 마커 정보와 사용자 유전형 정보를 이용하여 유전적 위험도를 계산하고, 임상설문 정보를 이용하여 환경요인별 증가 위험도를 검출한다(S32).

[0118] 아래 [표 3]을 참고하면, 흡연에 대한 증가 위험도는 1.8이고, 음주에 대한 증가 위험도는 1.5이다.

[0119] 이후, 수학식 GRS * (e_1*e_2*…*e_n) 또는 수학식 Pg*GRS + Pe*(e_1*e_2*…*e_n)을 이용하여 유전적 / 환경적 요인 통합 위험도를 산출한다(S34).

[0120] 이때, 유전적 위험도(GRS)는 유전적 요인(예를 들면, 위험인자 등)을 이용하여 다양한 방법으로 계산되는 값으로서, 본 발명에 따른 제 3 실시예에서는 유전적 위험도를 계산하는 방법을 한정하고 있진 않지만, 설명을 위해 제 1 실시예의 수학식 $\prod W_i * F_i$을 이용하여 계산하는 것을 예시로 한다. 따라서, 제 1 실시예에 따라 피검자 1의 GRS값은 (1*0.21)*(1*0.24)*(1*0.12)=0.006048이고, 피검자 2의 GRS값은 (1.66*0.21)*(1*0.24)*(1.85*0.12)=0.01857이고, 피검자 3의 GRS값은 (1*0.21)*(2.16*0.24)*(1*0.12)=0.01306이다.

[0121] 피검자들의 GRS값과 [표 3]에 나타난 값들을 수학식 GRS * (e_1*e_2*…*e_n)에 적용하여 계산해보면, 피검자 1의 유전적 / 환경적 요인 통합 위험도는 0.006048*(1.8*1.5) = 0.01633이고, 피검자 2의 유전적 / 환경적 요인 통합 위험도는 0.01857*(1*1) = 0.01857이고, 피검자 3의 유전적 / 환경적 요인 통합 위험도는 0.01306*(1*1.5) = 0.01960이다.

[0122] 여기서, 설문문항에 대해 사용자가 '아니오'라고 답한 경우, 요인별 증가 위험도는 1로 하여 계산한다.

[0123] 또한, 유전적 / 환경적 요인 통합 위험도는 수학식 Pg*GRS + Pe*(e_1*e_2*…*e_n)를 이용하여 산출될 수도 있다. 이때, Pg는 유전적 요인이 해당 질병 및 표현형에 미치는 비율이고, Pe는 환경적 요인이 해당 질병 및 표현형에 미치는 비율을 말한다.

[0124] 질병에 대해 보고된 유전적 요인이 10%, 환경적 요인이 90%라고 할 경우에, 수학식 Pg*GRS + Pe*(e_1*e_2*…*e_n)을 이용하여 유전적 / 환경적 요인 통합 위험도를 계산해보면, 피검자 1의 유전적 / 환경적 요인 통합 위험도는 0.1*0.006048 + 0.9*(1.8*1.5) = 2.4306이고, 피검자 2의 유전적 / 환경적 요인 통합 위험도는 0.1*0.01857 + 0.9*(1*1) = 0.9019이고, 피검자3의 유전적 / 환경적 요인 통합 위험도는 0.1*0.01306 + 0.9*(1*1.5) = 1.3513이다.

[0125] 여기서, 설문문항에 대해 사용자가 '아니오'라고 답한 경우, 요인별 증가 위험도는 1로 하여 계산한다.

표 3

[0127]

환경 요인	요인별 증가 위험도	피검자 1	피검자 2	피검자 3
흡연	1.8	네	아니오	아니오
음주	1.5	네	아니오	네

[0129] 본 발명은 또한 컴퓨터로 읽을 수 있는 기록 매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터로 읽을 수 있는 기록 매체는 컴퓨터에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록 장치를 포함한다. 컴퓨터가 읽을 수 있는 기록 매체의 예로는 롬(ROM), 램(RAM), 씨디-롬(CD-ROM), 자기 테이프, 플로피 디스크, 광 데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한, 컴퓨터가 읽을 수 있는 기록 매체는 유무선 통신망으로 연결된 컴퓨터 장치에 분산되어 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.

[0130] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위 내에서 다양한 수정, 변경, 및 치환이 가능할 것이다. 본 발명의 보호 범위는 아래의 청구 범위에 의해서 해석되어야 하며, 그와 동등한 범위

내에 있는 모든 기술 사상은 본 발명의 권리 범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

[0132] 100 : 위험도 예측 장치 120 : 획득부

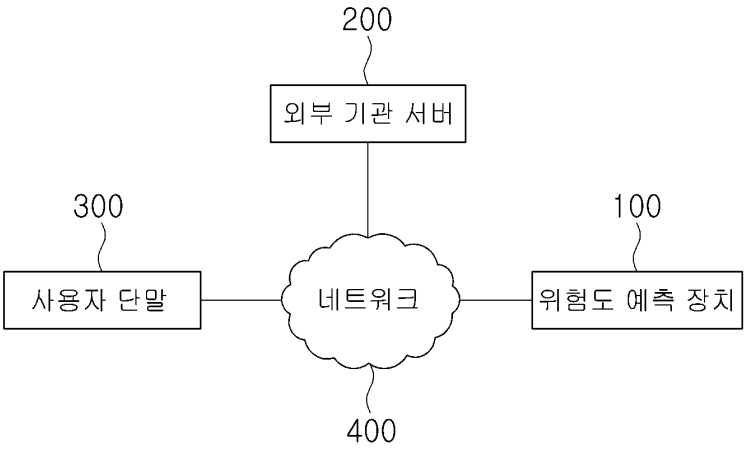
140 : 검출부 140 : 산출부

200 : 외부 기관 서버 300 : 사용자 단말

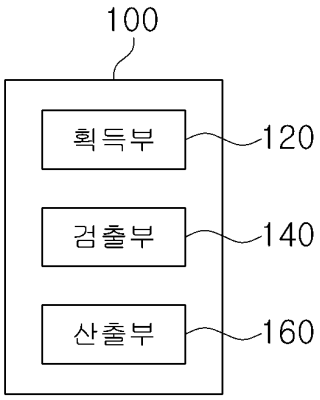
400 : 네트워크

도면

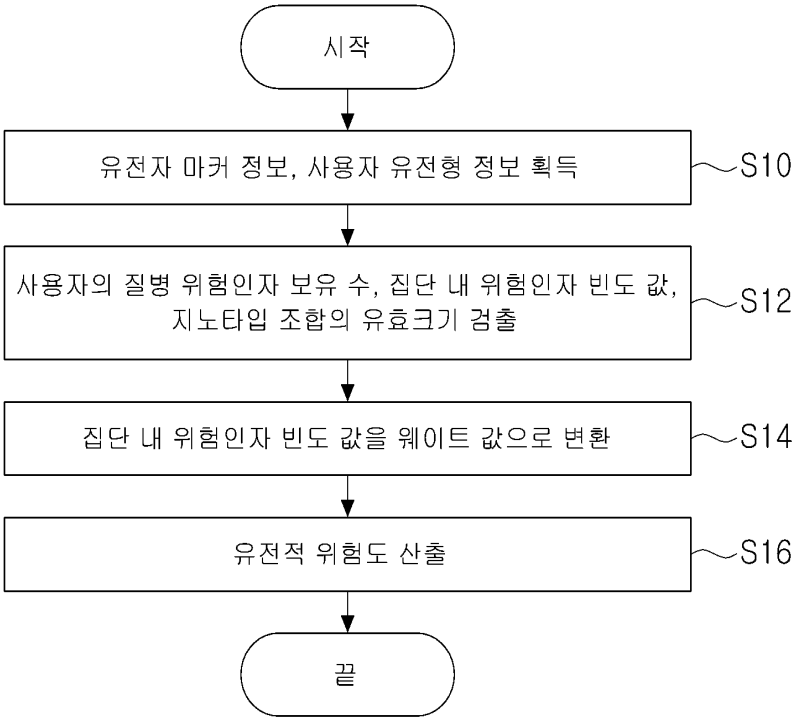
도면1



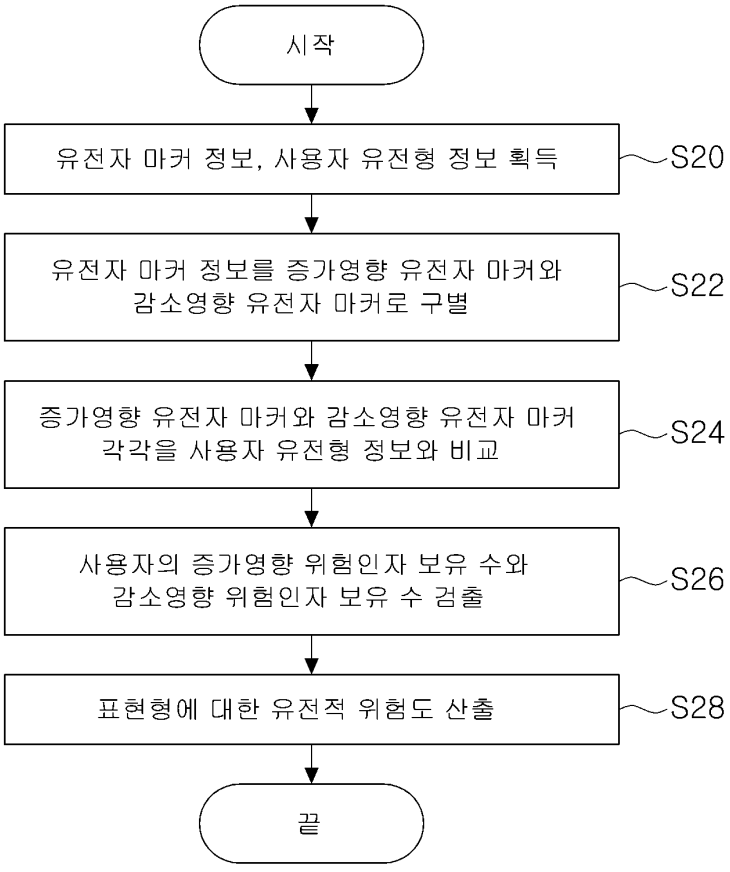
도면2



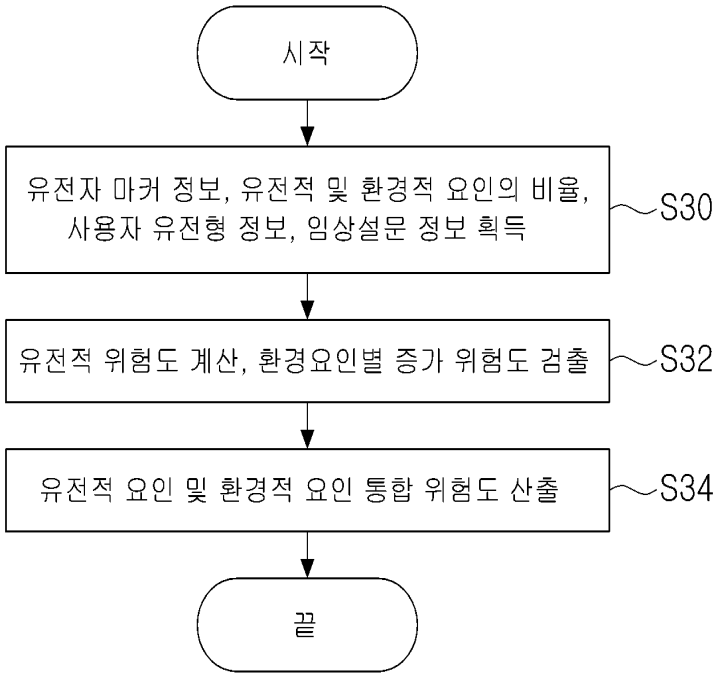
도면3


시작
유전자 마커 정보, 사용자 유전형 정보 획득
S10
사용자의 질병 위험인자 보유 수, 집단 내 위험인자 빈도 값, 지노타입 조합의 유효크기 검출
S12
집단 내 위험인자 빈도 값을 웨이트 값으로 변환
S14
유전적 위험도 산출
S16
끝

도면4


시작
유전자 마커 정보, 사용자 유전형 정보 획득
S20
유전자 마커 정보를 증가영향 유전자 마커와 감소영향 유전자 마커로 구별
S22
증가영향 유전자 마커와 감소영향 유전자 마커 각각을 사용자 유전형 정보와 비교
S24
사용자의 증가영향 위험인자 보유 수와 감소영향 위험인자 보유 수 검출
S26
표현형에 대한 유전적 위험도 산출
S28
끝

도면5


시작
유전자 마커 정보, 유전적 및 환경적 요인의 비율, 사용자 유전형 정보, 임상설문 정보 획득
S30
유전적 위험도 계산, 환경요인별 증가 위험도 검출
S32
유전적 요인 및 환경적 요인 통합 위험도 산출
S34
끝