



- (51) **International Patent Classification:**
G16H 50/70 (2018.01) *G16H 20/10* (2018.01)
- (21) **International Application Number:**
PCT/US2022/042262
- (22) **International Filing Date:**
31 August 2022 (31.08.2022)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
63/238,851 31 August 2021 (31.08.2021) US
63/250,912 30 September 2021 (30.09.2021) US
PCT/US2022/032250
03 June 2022 (03.06.2022) US
PCT/US2022/038941
29 July 2022 (29.07.2022) US
- (71) **Applicant:** **GUARDANT HEALTH, INC.**, [US/US]; 3100 Hanover St, Palo Alto, California 94304 (US).
- (72) **Inventors:** **KUMAR, Naveen**; 3100 Hanover St, Palo Alto, California 94304 (US). **ZHANG, Jingwen**; 3100 Hanover St, Palo Alto, California 94304 (US). **SUBRAMANIAN, Nisha**; 3100 Hanover St, Palo Alto, California 94304 (US). **NAYAK, Gautam**; 3100 Hanover Street, Palo Alto, California 94304 (US).
- (74) **Agent:** **HOTT, Timothy A.** et al.; 3100 Hanover St, Palo Alto, California 94304 (US).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE,

(54) **Title:** DATA REPOSITORY, SYSTEM, AND METHOD FOR COHORT SELECTION

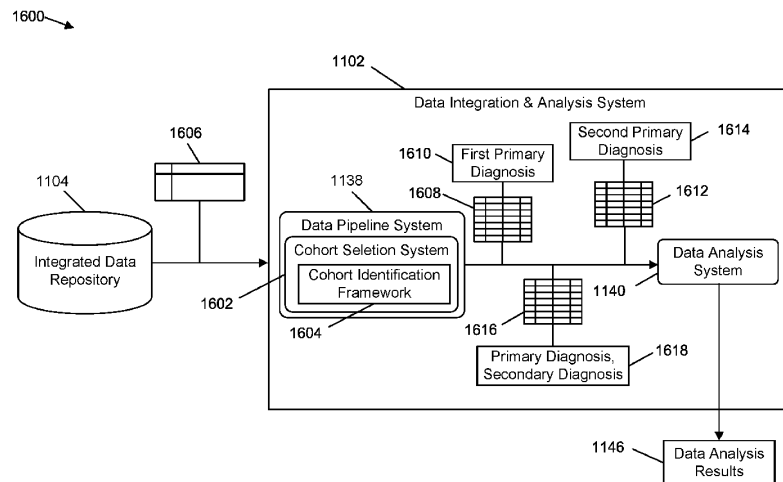


FIG. 16

(57) **Abstract:** A computer may access one or more medical data tables storing medical insurance transaction data for a plurality of patients. The one or more medical data tables comprise a date column and a diagnosis column. The computer may identify, based on the diagnosis column, a set of patients having a biological condition. The set of patients can be from among the plurality of patients. The computer may determine, for each patient in the set of patients, an earliest date when the patient received a diagnosis of the biological condition. The computer may identify, based on the diagnosis column and the date column, a cohort of patients from the set of patients. The cohort of patients can lack a diagnosis from a collection of biological conditions associated with a date occurring during a predefined time window before the earliest date when the patient received the diagnosis of the biological condition.

[Continued on next page]

KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*