

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
04 June 2020 (04.06.2020)



(10) International Publication Number
WO 2020/112478 A1

(51) International Patent Classification:
A61B 5/05 (2006.01) *A61B 5/117* (2016.01)
A61B 5/103 (2006.01)

(21) International Application Number:
PCT/US2019/062561

(22) International Filing Date:
21 November 2019 (21.11.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/773,028 29 November 2018 (29.11.2018) US
62/783,733 21 December 2018 (21.12.2018) US

(71) Applicant: **SOMALOGIC, INC.** [US/US]; 2945 Wilder-
ness Place, Boulder, Colorado 80301 (US).

(72) Inventors: **HAGAR, Yolanda**; 2821 20th Street, Boul-
der, Colorado 80304 (US). **DATTA, Gargi**; 2945 Wilder-
ness Place, Boulder, Colorado 80301 (US). **ALEXANDER,**
Leigh; 334 Pheasant Run, Louisville, Colorado 80027 (US).
HINTERBERG, Michael; 1773 Sphene Place, Loveland,
Colorado 80537 (US).

(74) Agent: **SANNY, Tony** et al.; 4940 Pearl East Circle, Suite
200, Boulder, Colorado 80301 (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, CZ, DE, DJ, DK, DM, DO,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,
HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, 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,

(54) Title: METHODS FOR DETERMINING DISEASE RISK COMBINING DOWNSAMPLING OF CLASS-IMBALANCED
SETS WITH SURVIVAL ANALYSIS

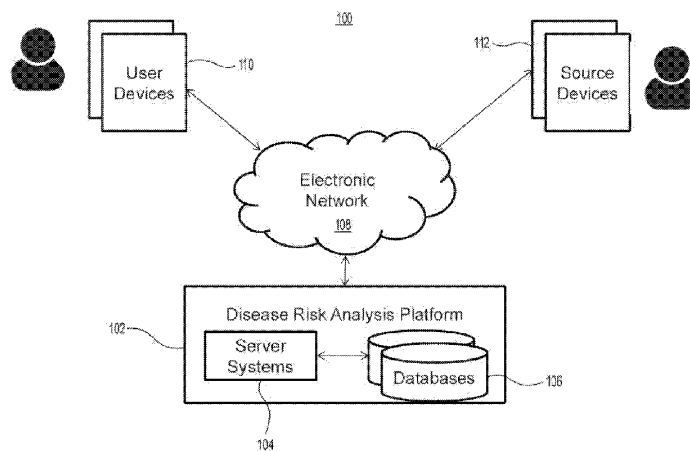


FIG. 1

(57) Abstract: A method for downsampling class-imbalanced sets with survival analysis comprising: acquiring a class-imbalanced data set, wherein the class-imbalanced data set comprises biological data from a plurality of subjects, wherein the biological data of each subject includes an observation, a time value, and a plurality of clinical measurements, and wherein the biological data is categorized as being part of a majority data class or a minority data class, wherein the majority data class has a greater number of observations than the minority data class; downsampling the class-imbalanced data set, wherein the downsampling results in the majority data class having an equivalent or substantially equivalent number of observations as the minority data class; and performing cross-validation on the downsampled data set with a survival analysis to generate a survival model, wherein the observation comprises an event or no event at a specific time value.

[Continued on next page]

SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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))