

METHOD FOR PREDICTING RISK OF DEVELOPING MYOPIA

TECHNICAL FIELD

5 [0001] The present disclosure relates to a method for predicting a risk of an individual to develop myopia.

BACKGROUND

[0002] Myopia, or short-sightedness, is the most common cause of vision loss worldwide, affecting approximately one-third of the global population. The prevalence of myopia is
10 increasing, with up to half of the world's population, or 5 billion people, predicted to have myopia by 2050. The increase in prevalence of myopia is accompanied by a decrease in the average age at which individuals are developing myopia, resulting in increased progression rate and myopia of greater severity when myopia stabilizes in teenagerhood or early adulthood.

[0003] Severe myopia, also known as high myopia, is associated with an increased risk of
15 developing sight-threatening conditions such as myopic macular degeneration, retinal detachments, glaucoma, and severe cataracts. As many as 1 billion people are predicted to have high myopia by the year 2050, placing a significant and potentially unmanageable proportion of the global population at risk of irreversible vision loss.

[0004] Myopia and high myopia represent significant challenges to the global community in
20 terms of personal and economic costs as well as loss of productivity, and there is an urgent need to develop new interventions to manage this epidemic. Research indicates that myopia may be prevented or slowed through non-medicinal interventions by managing its modifiable risk factors. A combination of outdoor activity and natural light exposure, looking far into the distance, taking regular breaks from near-work activities such as reading and using smart
25 devices, and undergoing annual eye check-ups during childhood have been associated with slowing or halting of myopia onset and progression. However, such strategies are preventive and require early intervention to be effective. Therefore, knowing the risk of an individual to develop myopia may assist in such early intervention to maximise the individual's chances of having healthy vision.

[0005] Existing systems and methods to predict a risk of developing myopia have limited input factors, and outputs generated are of limited value or benefit to a subject. A subject would thus be unable to use such outputs generated to intervene and prevent the development of myopia. There is therefore a need for a more accurate tool to predict risk of developing myopia, which
5 further provides preventive and prescriptive outcomes.

SUMMARY

[0006] One general aspect of the present disclosure includes a computer-implemented method for predicting a risk of developing myopia in a subject. The method comprises at a computer system having one or more computer processors, and a memory storing at least one program
10 for execution by the one or more computer processors, receiving, by the one or more computer processors via a computerized device, a subject data on a subject, wherein the subject data comprises a subject current age, a subject gender, a subject country of residence, a subject eye test compliance data, a subject screen time data, a subject outdoor data, a subject parental myopia status, accessing, by the one or more computer processors, at least one population data
15 array generated based on a population data, the population data comprising data received from different subjects and corresponding instances of myopia that occurred for the different subjects, said population data array is stored in at least one database, retrieving by the one or more computer processors, at least one population data array entry from the at least one population data array, generating, by the one or more computer processors, at least one variable
20 based on at least the received subject data, determining the risk of the subject to develop myopia, by applying at least one model, by the one or more processors, on the received subject data, the at least one population data entry and the generated at least one variable, said model is trained based on individual data arrays generated based on the population data, generating a score indicating a likelihood of the subject to develop myopia, based on the determined risk;
25 and generating a report comprising the score indicating the likelihood of the subject to develop myopia.

[0007] Implementations may include one or more of the following features. The method further comprises modifying the score, by the one or more computer processors, based on updated received subject data and updated received population data array over time. The method further
30 comprises storing the risk of the subject to develop myopia. The step of generating the at least one variable in the method comprises: generating at least one individual data array comprising at least one data point for each different subject; generating a plurality of decision trees based

on the at least one individual data array and identifying a plurality of decision tree nodes; identifying at least one interaction between the plurality of decision trees; and selecting at least one variable from the at least one data point, the plurality of decision tree nodes, and the at least one interaction, based on significance. The at least one population data array comprises a

5 first population data array, the first population data array comprising at least one age and a mean break session data associated with each of the at least one age; and the at least one population data array entry comprises a first population data array entry, the first population data array entry comprising the mean break session data associated with the at least one age corresponding to the subject current age. The at least one population data array comprises a

10 second population data array, the second population data array comprising at least one age and a mean break frequency data associated with each of the at least one age; and the at least one population data array entry comprises a second population data array entry, the second population data array entry comprising the mean break frequency data associated with the at least one age corresponding to the subject current age. The at least one variable comprises a

15 first variable, the first variable comprising a subject eye test age based on the subject current age and the subject eye test compliance data. The at least one population data array comprises a third population data array, the third population data array comprising at least one age and a proportion of population with myopia onset associated with each of the at least one age. The at least one population data array entry comprises a third population data array entry, the third

20 population data array entry comprising an at least one age associated with the highest proportion of population with myopia onset. The at least one variable comprises a second variable based on a chronical age at which myopia first appeared, the second variable generated based on the third population data array entry and a subject eye test age. The at least one variable is selected from the group comprising of a third variable, the third variable based on

25 whether the second variable is more than or is equal to 0; and a fourth variable, the fourth variable based on whether the second variable is more than or is equal to 6. The at least one population data array further comprises a fourth population data array, the fourth population data array comprising at least one age and a mean eye test age associated with each of the at least one age; and the at least one population data array entry further comprises a fifth

30 population data array entry, the fifth population data array entry comprising a mean eye test age associated with the at least one age corresponding to the subject current age. The at least one variable comprises a fifth variable based on a fourth population data array entry and the second variable. The at least one population data array entry comprises a fifth population data array entry, the fifth population data array entry comprising a proportion of population with

myopia onset associated with an at least one age corresponding to the subject current age. The at least one variable is selected from the group comprising of: a sixth variable based on whether a fifth variable is more than or equal to 0.012; a seventh variable based on whether the fifth variable is less than 0.075; an eighth variable based on whether the fifth variable is less than 0.098; and a ninth variable based on whether the fifth variable is less than 0.045. The at least one variable comprises a tenth variable, the tenth variable based on whether a differential between the current age and a first variable is more than 8. The at least one population data array comprises a fifth population data array, the fifth population data array comprising at least one age and a mean eye test age difference associated with each of the at least one age, the mean eye test age difference based on a differential between a user current age and an eye test age of the population data; and the at least one variable comprises an eleventh variable, the eleventh variable based on whether a fifth population data array entry is more than or equal to 6.7. The at least one variable comprises a twelfth variable, the twelfth variable based on a fifth population data array entry and a differential between the subject current age and a subject eye test age. The at least one variable comprises a thirteenth variable, the thirteenth variable based on the subject screen time data. The method further comprises applying at least one risk ratio based on the subject parental myopia status. The method further comprises applying at least one risk ratio based on the subject outdoor data. The method further comprises applying at least one risk ratio based on the subject gender. The method further comprises applying at least one risk ratio based on the subject screen time data. The method further comprises applying at least one risk ratio based on the subject eye test compliance data. The generated report comprises recommendations to a user on actions to reduce risk of the subject developing myopia.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] Some non-limiting exemplary embodiments or features of the disclosed subject matter are illustrated in the following drawings.

In the drawings:

[0009] Fig. 1 is a high-level diagram schematic of a computer system for generating and modelling predictions of a risk of developing myopia, in accordance with some embodiments of the present disclosure;

- [0010] Fig. 2 is a schematic illustration of a computer-implemented prediction module 200 for generating predictions of a risk of developing myopia, in accordance with embodiments of the present disclosure;
- 5 [0011] Fig. 3 is a schematic illustration of example inputs for at least one model for predicting risk of a subject developing myopia, in accordance with embodiments of the present disclosure;
- [0012] Fig. 4A is a schematic illustration of a Ranger survival model plot illustrating a relationship between survival probability of myopia incidence over time, in accordance with embodiments of the present disclosure;
- 10 [0013] Fig. 4B is a schematic illustration of a Random Forest Ranger model plot illustrating a relationship between the average survival probability of myopia incidence over time, per all of the parameters implemented in the Random Forest Ranger model, in accordance with embodiments of the present disclosure;
- [0014] Fig. 5 is a schematic illustration of a method of predicting a risk of developing myopia, in accordance with embodiments of the present disclosure.
- 15 [0015] With specific reference now to the drawings in detail, it is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of embodiments of the disclosure. In this regard, the description taken with the drawings makes apparent to those skilled in the art how embodiments of the disclosure may be practiced.
- 20 [0016] Identical or duplicate or equivalent or similar structures, elements, or parts that appear in one or more drawings are generally labeled with the same reference numeral, optionally with an additional letter or letters to distinguish between similar entities or variants of entities, and may not be repeatedly labeled and/or described. References to previously presented elements are implied without necessarily further citing the drawing or description in which they appear.
- 25 [0017] Dimensions of components and features shown in the figures are chosen for convenience or clarity of presentation and are not necessarily shown to scale or true perspective. For convenience or clarity, some elements or structures are not shown or shown only partially and/or with different perspective or from different point of views.

DETAILED DESCRIPTION

[0018] In the following detailed description, numerous specific details are set forth in order to provide a thorough understanding of the invention. However, it will be understood by those of ordinary skill in the art that the invention may be practiced without these specific details. In other instances, well-known methods, procedures, components, modules, units and/or circuits have not been described in detail so as not to obscure the invention.

[0019] Dimensions of components and features shown in the figures are chosen for convenience or clarity of presentation and are not necessarily shown to scale or true perspective. For convenience or clarity, some elements or structures are not shown or shown only partially and/or with different perspective or from different point of views.

[0020] Although embodiments of the disclosure are not limited in this regard, discussions utilizing terms such as, for example, “processing,” “computing,” “calculating,” “determining,” “establishing,” “analyzing,” “checking”, or the like, may refer to operation(s) and/or process(es) of a computer, a computing platform, a computing system, a mobile operating system or other electronic computing device, that manipulates and/or transforms data represented as physical (e.g., electronic) quantities within the computer’s registers and/or memories into other data similarly represented as physical quantities within the computer’s or smartphone’s registers and/or memories or other information non-transitory storage medium (e.g., a memory) that may store instructions to perform operations and/or processes.

[0021] Although embodiments of the invention are not limited in this regard, the terms “plurality” and “a plurality” as used herein may include, for example, “multiple” or “two or more”. The terms “plurality” or “a plurality” may be used throughout the specification to describe two or more components, devices, elements, units, parameters, or the like. Unless explicitly stated, the method embodiments described herein are not constrained to a particular order or sequence. Additionally, some of the described method embodiments or elements thereof can occur or be performed simultaneously, at the same point in time, or concurrently. Unless otherwise indicated, use of the conjunction “or” as used herein is to be understood as inclusive (any or all of the stated options).

[0022] With specific reference now to the drawings in detail, it is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of embodiments of

the disclosure. In this regard, the description taken with the drawings makes apparent to those skilled in the art how embodiments of the disclosure may be practiced.

[0023] Identical or duplicate or equivalent or similar structures, elements, or parts that appear in one or more drawings are generally labelled with the same reference numeral, optionally with an additional letter or letters to distinguish between similar entities or variants of entities and may not be repeatedly labelled and/or described. References to previously presented elements are implied without necessarily further citing the drawing or description in which they appear.

[0024] The present disclosure provides a computer method for reliably predicting the likelihood that a child (e.g., between the ages of 4 and 14) will develop myopia. Future computer methods will provide additional details such as at what age will myopia be developed, to what extent (e.g., expected rate of annual progression of myopia), whether the child will develop high myopia, at what age and degree (or dioptre) their myopic deterioration is likely to stabilise, and the cut-off age after which a child is highly unlikely to develop myopia.

[0025] The method of the present disclosure is an easy-to-use web-based platform that utilizes inputs such as personal information concerning age, gender, country of residence, whether an eye test was completed, parental myopia status, as well as behaviour related information, such as screen time per day, and outdoor time, thereby providing inputs that may be used to model the future changes in vision, with state-of-the-art machine learning technology, Artificial intelligence (AI), and big data analytics to reliably predict the likelihood that any specific child will develop myopia in the future. That is, the prediction is personalized per user and is provided in real-time with no need for acquiring any biological specimen from the user. Currently, there is a lack of prescriptive information available to parents that explains how they may reduce their child's risk of myopia. The method of the present disclosure may provide parents with recommendations, based on recent and most updated scientific evidence, to manage their child's individual risk factors so that they may maximize their chances of having healthy vision now and in the future.

[0026] Data that is collected from a current userbase, as well as data that will be collected from an expectedly growing userbase will enable to accrue a large database on the risk factors, prevalence, incidence and progression of myopia in an unprecedentedly large and diverse population, which will be used to answer a variety of important research questions. As more

user data are collected, additional information will be inputted into the predictive algorithm or predictive model of the present disclosure, to improve its accuracy. As data will continue to be inputted and collected, the predictive model will in turn be changed to be improved, if necessary, based on machine learning and AI and big data analytics. By evaluating the predictions made by the prediction model based on the specific risk factor data supplied by parents as well as real-time data collected by computerized devices, with the addition of data collected from scientific literature, it will be possible to ascertain the true influence of each of these risk factors longitudinally on the onset and progression of myopia.

[0027] Fig. 1 is a high-level diagram schematic of a computer system 100 for generating and modelling predictions of a risk of developing myopia, in accordance with some embodiments of the present disclosure. Computer system 100 may be any computer system or device, including desktop computers and portable electronic devices such as smartphones, tablet computers and laptops. In some embodiments, computer system 100 may comprise one or more computer processors 102, and at least one memory 104. The one or more computer processors 102 may include at least one processing unit and may be configured to execute programmed computer-executable instructions that are stored in the at least one memory 104. In some embodiments, the one or more computer processors 102 may run one or more modules, such as a prediction module 122, population data array module 124 and data analysis module 126. In some embodiments, the prediction module 122 may interact with population data array module 124 and data analysis module 126. In some embodiments, prediction module 122 may be used to predict a risk of developing myopia of a subject. In some embodiments, population data array module 124 may be used to generate at least one population data array, which is an array of data relating to various parameters collected from a population, e.g., a population of subjects. Population data array module 124 may further be used to generate at least one population data array entry for retrieval by the prediction module 122 and/or the data analysis module 126. In some embodiments, data analysis module 126 may be used to analyse a population data and/or data input and generate output for the prediction module 122.

[0028] According to some embodiments, the at least one memory 104 may be a volatile or non-volatile memory and may comprise data storage 106 and program memory 108. In some embodiments, data storage 106 may store an executable version of a software useful to practice the techniques and methods described in the present disclosure. In some embodiments, data storage 106 may be included in the at least one memory 104 or may be separate from the at

least one memory 104. In some embodiments, computer system 100 may comprise a volatile memory and a non-volatile memory. In some embodiments, the at least one memory 104 may interact with the one or more computer processors 102. In some embodiments, the at least one memory 104 may be used to store an executable version of a software, or computer application
5 useful to practice the techniques and methods described in the present disclosure.

[0029] According to some embodiments, computer system 100 may further comprise an I/O interface 110 to transfer information between computer system 100 and external devices such as a display 132, a keyboard 134 and/or a selection device 136. Display 132 may generally operate to provide a presentation of content. In some embodiments, display 132 may be any
10 computer monitor or screen. In some embodiments, one or more computer processors 102 may communicate with display 132 to project the content, e.g., a report comprising a score indicating prediction of developing myopia, on the display 132. In some embodiments, selection device 136 may be any input or pointing device for a user to input information, data, or instructions for operation of computer system 100, including a mouse or a touch display
15 integrated with display 132. In some embodiments, computer system 100 may further comprise a network interface or adapter 112 to connect to a local network or a wireless network.

[0030] It should be recognized that the software functionality described in Fig. 1 is an example and that the functions described herein may be implemented in hardware, software, firmware, or any combination thereof. If implemented in software, the functions may be stored on, or
20 transmitted as, one or more instructions or code encoded on a non-transitory computer-readable medium, such as storage 106. According to some embodiments, the methods described within the present disclosure may be stored as instructions on the computer-readable medium to await execution by a computer system, e.g., computer system 100.

[0031] Fig. 2 is a schematic illustration of a computer-implemented prediction module 200 for
25 generating predictions of a risk of developing myopia, in accordance with embodiments of the present disclosure. In some embodiments, prediction module 200, may be hardware or a combination of hardware and software, similar to prediction module 122 (Fig. 1). Predictions regarding risk of developing myopia may be generated by at least one model 240 based on subject data 236 of a specific subject for whom prediction of risk of developing myopia is
30 sought for. In some embodiments, subject data 236 may be received by a computer system, e.g., computer system 100. In some embodiments, subject data 236 may be subject related information, which may be inputted by a user, such as the subject themselves or a healthcare

provider, to obtain a prediction of a risk of the subject developing myopia. In some embodiments, as myopia is a disease that generally onsets when the subject is young and stabilizes in the late teens, the user may be a parent, supervisor or guardian of the subject, e.g., a child, who is interested in determining a risk of the subject developing myopia. In some
5 embodiments, the user may use a computerized user device, such as a personal computer or smartphone to interact with module 200, input subject data 236 and instruct module 200 to analyse subject data 236, generate reports 250, and perform other actions. In some embodiments, any information acquired, processed, and generated by module 200 may be stored on one or more storage devices, e.g., data storage 106 (Fig. 1). In some embodiments,
10 the storage device may be network-based.

[0032] In some embodiments, module 200 may receive subject data 236 from a user, apply the at least one stored model 240 to the subject data 236, and provide results of application of the at least one model 240 as at least one report 250 to the user. In some embodiments, the results may be provided in electronic form on a web-based form or on a graphical user interface to be
15 displayed on a display of the user device, e.g., display 132 (Fig. 1). In some embodiments, the at least one report 250 may comprise an indication of the risk of the subject developing myopia. In some embodiments, the at least one report 250 may comprise a score indicating the likelihood of the subject to develop myopia. The score may be based on the risk of the subject to develop myopia, whereby the risk may be determined by applying at least one model, e.g.,
20 at least one model 240, by the one or more processors, on the received subject data 236, the retrieved at least one population data entry and the generated at least one variable (as detailed hereinbelow). In some embodiments, the score may be represented by a number, such as a percentage of likelihood of a subject to develop myopia at a certain age. In other embodiments, the score may be represented by some sort of gauge indicating low, medium or high levels of
25 likelihood of the subject to develop myopia at a certain age. It should be clear that other score representations may be used.

[0033] In some embodiments, the at least one report 250, may include one or more recommendations to a user on one or more actions or behavioural changes that may be made by the subject to potentially reduce a risk of developing myopia. For example, a
30 recommendation to a user may include spending less time in front of a computerized device, spending more time outdoors, spending more time in the sun (during daytime), intaking Vitamins, such as Vitamin D, going to a check-up within a certain time period, intaking

Atropine eye drops, using contact lenses, wearing prescriptive spectacles, performing orthokeratology with specially made contact lenses, and so on.

[0034] According to some embodiments of the present disclosure, prediction by the at least one model 240 may also be based on at least one generated variable 232 based on subject data 236 and information stored in at least one database 230, stored on one or more storage devices associated with system 100. In some embodiments, at least one database 230 may be generated by at least one data module 224 of module 200. In some embodiments, the at least one data module 224 may receive, generate, process, or analyse a collection of population data 226 collected from a population of subjects whose information is present in the module 200. For some information, the population data 226 may be stored in at least one database 230.

[0035] According to some embodiments of the present disclosure, the population data 226 may comprise raw data 202 from different subjects comprising the population of subjects and corresponding instances of myopia that occurred for each of the different subjects. In some embodiments, the raw data 202 of each subject within the population may comprise characteristics data 204, behaviour data 206 and myopic condition data 208 of each of the different subjects to be stored in at least one database 230. Characteristic data 204 may include data on demographic, geographic, or physical characteristics. In some embodiments, characteristic data 204 may comprise data on each subject's age (e.g., between 4 to 14), gender (male or female), parental myopia status (e.g., none, one parent or both parents suffering from myopia), eye test age (the age at which a subject had gone for an eye check-up), country of residence, number of siblings, and sibling myopia status. In some embodiments, behaviour data 206 may comprise data received in real-time from the user, e.g., via a computerized user device, such as a personal computer or smartphone, which may communicate with module 200. In some embodiments, behaviour data 206 may comprise data on each subject's behavioural habits, including the average number of hours spent by the subject on outdoor activities per day (e.g., 0-20 hours per week), an average time spent by each subject viewing electronic screens or displays, such as computers or smart phones (e.g., 0-12 hours per day), a frequency of break sessions from viewing electronic screens or displays taken by each subject per day, a duration of break sessions from viewing electronic screens or displays taken by each subject per day, and a number of indications provided to each subject indicating improper screen viewing environments, such as near face-to-screen distance alerts, poor posture alerts, low ambient light alerts, and spectacle detection alerts. In some embodiments, myopic condition

data 208 may comprise data on each subject's myopic condition (if any) and may comprise data such as an age of myopia onset, the age at which the subject commenced wearing spectacles (e.g., 2-16 years old), a severity of myopia, a degree of myopia of each eye, and the subject's degree of myopia at the time of commencement of spectacle use for both left and right eyes (e.g., 0-1000). In some embodiments, the raw data 202 may be based on input received from the population subjects. In some embodiments, additional and/or other parameters may be inputted by the user.

[0036] In some embodiments, the raw data 202 may also be generated based on usage of a software, application, or sensor configured to receive information about each subject's usage of devices with electronic screens or displays. In some embodiments, raw data 202 may be generated using an application installed on a smart phone or computer of each subject. For example, at least some of the raw data 202 may be automatically generated using the plano®App, by Plano Pte Ltd., which provides parental control for managing a child's device use and screen time. In some embodiments, the raw data 202 may be collected over a period of time to analyse any changes in behaviour data 206 and myopic condition data 208 over time and to determine the impact of various data types on the development of myopia.

[0037] According to some embodiments of the present disclosure, raw data 202 may be assigned values or representations. Some examples of raw data 202 and their assigned values or representations include: CurrentAge (assigned a value equal to the subject's age), MyoGroup (assigned a binary representation identifying whether the subject has myopia in both eyes), EyeTest (assigned a binary representation identifying whether the subject has taken an eye test), EyeTestAge (assigned a value equal to the age in years which the subject took an eye test), EyeTestComp (assigned a binary representation representing whether the subject has taken an eye test within the past year), OnsetAge (assigned a value equal to the subject's age when myopia developed), Country (assigned a value based on the country of residence of the subject), Gender (assigned a value based on the gender of the subject), RightEye (assigned a value equal to the myopic degree of the right eye), LeftEye (assigned a value to the myopic degree of the left eye), StartedRightEye (assigned a value to a myopic degree of the right eye before development of myopia), StartedLeftEye (assigned a value to a myopic degree of the left eye before development of myopia) and AgeGroup (assigned a value based on the age group that the subject belongs to, e.g., {2-5, 6-10, 11-16}).

[0038] In some embodiments, EyeTestComp that may be included in raw data 202, may be used as a reference to prove that those subjects that are eye test compliant, i.e., who have good habits of having annual eye checks, have a lower risk of transiting from Non-myopic to Myopic (e.g., these subjects show a risk ratio of 0.9989, $p < 0.05$).

5 [0039] According to some embodiments, the at least one data module 224 may generate at least one feature 210 for each subject based on the raw data 202. Features 210 may include data points calculated based on raw data 202. In some embodiments, features 210 may be assigned values or representations. Examples of features 210 and their assigned values or representations include: EyeTestDiff (assigned a value equal to a difference in years between CurrentAge and
10 EyeTestAge), TimeDiff (assigned a value equal to a difference between CurrentAge and OnsetAge), ProgRate (assigned a value equal to an annual progression rate of myopia calculated based on a difference in myopia severity divided by TimeDiff), screen_time_hr (assigned a value based on a mean number of hours per day spent by the subject viewing electronic screens or displays), Incidence (onset of the disease), BreakSess (assigned a value
15 corresponding to a mean duration of breaks from viewing electronic screens or displays taken by the subject), and BreakFreq (assigned a value corresponding to a mean number of breaks from viewing electronic screens or displays taken by the subject).

[0040] According to some embodiments, the at least one data module 224 may further generate and store at least one population data array 212 based on population data 226 which comprises
20 the raw data 202 and features 210 for all the different subjects. In some embodiments, the at least one population data array 212 may be in the form of lookup tables. In some embodiments, different population data arrays 212 may be generated per a specific parameter, for example, different population data arrays may be generated for different countries of residence. The at least one population data array 212 may comprise at least one age, and at least one data
25 associated with each of the at least one age.

[0041] According to some embodiments of the present disclosure, population data array 212 may comprise a first population data array comprising at least one age and a mean break session data associated with each of the at least one age (breakSessAgeTab). An example of first population data array is illustrated as Table 1.

Table 1: Mean break sessions per day by current age

Current Age	Mean Break Sessions Per Day (breakSessAgeTab)
2	118.53
3	86.39
4	60.15
5	2328.88
6	3594.11
7	1169.9
8	12365.55
9	8094.14
10	834.52
11	628.4
12	3096.77
13	306.32
14	554.32

[0042] According to some embodiments of the present disclosure, population data array 212 may comprise a second population data array comprising at least one age and a mean break frequency data associated with each of the at least one age (breakFreqAgeTab). An example of second population data array is illustrated as Table 2.

Table 2: Mean break frequency per day by current age

Current Age	Mean Break Frequency Per Day (breakFreqAgeTab)
2	1.23
3	0.94
4	0.65
5	0.83
6	1.06
7	0.96
8	1.00

9	1.06
10	1.16
11	1.21
12	1.16
13	1.09
14	1.15
15	0.88
16	1.45

[0043] According to some embodiments of the present disclosure, population data array 212 may comprise a third population data array comprising at least one age and a proportion of population with myopia onset associated with each of the at least one age (OnsetAgeTab). An example of third population data array is illustrated as Table 3.

Table 3: Proportion of population with myopia onset by current age

Current Age	Population Proportion (OnsetAgeTab)
2	0.070145556
3	0.038866522
4	0.058996593
5	0.083462372
6	0.115670486
7	0.137503871
8	0.17698978
9	0.1196965
10	0.112263859
11	0.04227315
12	0.032208114
13	0.006968102
14	0.003716321
15	0.001238774

[0044] According to some embodiments of the present disclosure, population data array 212 may comprise a fourth population data array comprising at least one age and a mean eye test age associated with each of the at least one age (EyeTestTab). An example of fourth population data array is illustrated as Table 4.

5

Table 4: Mean eye test age by current age

Current Age	Mean Eye Test Age (EyeTestTab)
2	0
3	0.03
4	0.13
5	0.07
6	0.15
7	0.27
8	0.47
9	0.9
10	1.26
11	1.97
12	2.45
13	2.96
14	2.95
15	2.84
16	2.93

[0045] According to some embodiments of the present disclosure, population data array 212 may comprise a fifth population data array comprising at least one age and a mean eye test age difference associated with each of the at least one age, the mean eye test age difference based on a differential between the current age of each subject and the eye test age of each subject (EyeTestDiffTab). An example of fifth population data array is illustrated as Table 5.

10

Table 5: Mean eye test age difference by current age

Current Age	Mean Eye Test Age Difference (EyeTestDiffTab)
2	2
3	2.97
4	3.87
5	4.93
6	5.85
7	6.73
8	7.53
9	8.1
10	8.74
11	9.03
12	9.55
13	10.04
14	11.05
15	12.16
16	13.07

[0046] There are many population data arrays that may be generated, and any generated population data arrays disclosed, described, and discussed in this present disclosure are merely representative and should not be construed as a complete listing of population data arrays that may be generated.

[0047] According to some embodiments of the present disclosure, once the at least one population data array 212 has been generated, system 200 may assign each subject with at least one additional data point comprising a value representing a retrieved population data array entry from the at least one population data array 212. In some embodiments, the one or more processors, e.g., processor 102 (Fig. 1) may access the database in which the population data array is stored, and retrieve at least one population data array entry.

[0048] For example, a subject with current age 8 would have a value of 12365.55 as the retrieved population data array entry for breakSessAgeTab, a value of 1.00 as the retrieved

population data entry for breakFreqAgeTab, a value of 0.17698978 as the retrieved population data entry for OnsetAgeTab, a value of 0.47 as the retrieved population data entry for EyeTestTab, and a value of 7.53 as the retrieved population data entry for EyeTestDiffTab. In some embodiments, module 200 may then generate an initial individual data array 214 for each subject, the initial individual data array 214 comprising data points of the raw data 202 and features 210 of each subject, as well as the at least one retrieved population data array entry corresponding to the current age of the subject.

[0049] According to some embodiments of the present disclosure, module 200 may derive at least one derived variable 216 from the initial individual data arrays 214. In some embodiments, the at least one derived variable 216 may be derived by applying one or more algorithms or calculations to the initial individual data arrays 214 and selecting at least one derived variable. In some embodiments, the at least one derived variable 216 may be derived based on the data within the initial individual data arrays 214 and the at least one population data array 212. For example, there may be a derived variable 216 (e.g., MaxDiff2) which is based on an age associated with the highest proportion of population with myopia onset (i.e., chronological age at which myopia first appeared) indicated in the population data array 212 titled OnsetAgeTab, and the subject eye test age. For example, a subject with current age 9 and eye test age 0 (i.e., never had an eye test), would have a value of $9 - 0 = 9$ for the derived variable 216 named MaxDiff2. This value would then be added to the initial individual data array 214.

[0050] In some embodiments, the at least one derived variable 216 may further be derived by generating a plurality of decision trees based on the at least one initial individual data array 214 and identifying a plurality of decision tree nodes. In some embodiments, the plurality of decision trees may be generated to classify the subjects of the population into subjects with myopia and without myopia. In some embodiments, the plurality of decision trees may be generated with the raw data 202 data point MyoGroup (0 for subjects with no myopia and 1 for subjects suffering from myopia), ProgRate (progression rate of myopia), or OnsetAge (chronological age at which myopia first appeared) as the predictor. Examples of the at least one derived variable 216 generated, include whether the derived variable 216 named MaxDiff2 is more than or equal to 0 (dTree1, representing 'decision tree 1'), whether the derived variable 216 named MaxDiff2 is more than or equal to 6 (dTree4), whether the retrieved population data array entry from the population data array 212 titled OnsetAgeTab is more than or equal to 0.012 (dTree5), whether the retrieved population data array entry from the population data

array 212 titled OnsetAgeTab is less than 0.075 (dTree7), whether the retrieved population data array entry from the population data array 212 titled OnsetAgeTab is less than 0.098 (dTree9), whether the retrieved population data array entry from the population data array 212 titled OnsetAgeTab is less than 0.045 (dTree12), whether the retrieved population data array entry
5 from the population data array 212 titled EyeTestDiffTab is more than or equal to 6.7 (dTree15), and whether the value for the data point for EyeTestDiff is more than 8 (dTree17). In some embodiments, there may be additional decision trees that may be included as part of the prediction model. Typically, the values of the decision trees are either '0' or '1'. That is, if the condition (e.g., any of the conditions mentioned hereinabove per each tree, in parenthesis)
10 is met, per each decision tree, the value is '1' and if the condition is not met, the value will be '0'.

[0051] In some embodiments, the at least one derived variable 216 may further be derived by interactions between variables relating to a plurality of decision trees, based on the at least one initial individual data array 214 and identifying a plurality of decision tree nodes. That is, some
15 of variables 216 may be created by interactions between one or more variables, thereby creating additional 'interaction-based' variables. For example, the MyoGroup predictor may be created by the calculation of EyetestTab multiplied by MaxDiff2, whereas the ProgRate predictor may be created from EyetestDiff and EyeTestTab (i.e., the value of the interaction-based variable is '1' if both conditions of each of these variables EyetestDiff and EyeTestTab are met, and the
20 value of this interaction-based variable is 0 if at least one of the conditions of either of EyetestDiff and EyeTestTab is not met). In some embodiments, the decision trees show which input variables influence the predictor, e.g., MayGroup predictor, ProgRate predictor or OnsetAge predictor. According to some embodiments, a decision tree may output a set of rules, which are coded as '1' when a condition is met and '0' when a condition is not met. Based on
25 this output, variables may be created based on a sequence of interactions between the variables. Thus, in some embodiments, an interaction between variables may be a multiplication of at least two variables, e.g., at least two decision tree variables, whereas in other embodiments, an interaction between variables may be the addition of at least two variables, e.g., that the condition of each of the at least two variables is met, and not just of one of the at least two
30 variables. Other interactions between variables may also be implemented to thereby create additional variables 216.

[0052] According to some embodiments of the present disclosure, the at least one data module 224 may further generate and store information on one or more derived variables 216 in at least one database 230. The one or more derived variables 216 may be derived using one or more machine learning approaches based at least in part on combinations or interactions between the population data 226 and the at least one population data array 212. In some embodiments, the one or more derived variables 216 may be derived using one or more algorithms based at least in part on the characteristics 204, behaviour 206 and myopic condition 208 information. In some embodiments, data for each subject corresponding to the one or more derived variables 216 may be generated and stored. There are many variables that may be derived from the raw data 202 by using various algorithms to capture as much of the trends and patterns in the data that should be fed into the at least one model 240. It should be clear that each model of the at least one model 240, may use a subset of a pool of generated variables, based on how well these variables fit into that specific model, and how significant these variables are in improving the overall accuracy of the model. Any derived variables disclosed, described, and discussed in this present disclosure are merely representative and should not be construed as a complete listing of variables that may be derived.

[0053] According to some embodiments, prediction by the at least one model 240 may be based on at least one meta-analysis parameter 238 generated based on subject data 236. In other embodiments, the at least one meta-analysis parameter 238 may be generated based on one or more external data sources 237, such as literature and published documents. There are many external sources that may be used to generate several meta-analysis parameters, and any external sources or meta-analysis parameters disclosed, described, and discussed in this present disclosure are merely representative and should not be construed as a complete listing of external sources and meta-analysis parameters.

[0054] According to some embodiments, a meta-analysis that may generate at least one meta-analysis parameter 238 is a statistical analysis that attempts to integrate the results of several individual studies. There may be two types of meta-analysis models, namely the fixed effects and random effects model. The fixed effects model assumes that all studies measure the same common treatment effect (in this case myopia). The random effects model assumes that the effects of individual studies deviate from the true intervention effect of all studies due to sampling error, but that there is another source of variance introduced due to the fact that the studies do not stem from one single population but are drawn from a “universe” of

populations. According to the present disclosure, in order to calculate the overall effect of the studies, all effect sizes are averaged, while giving a higher weight to studies with high precision.

5 [0055] In the present disclosure, the meta-analysis model that may generate the at least one meta-analysis parameter 238, may be based on the effects of gender, parental myopia status, outdoor activity, and screen time, which were obtained from literature. Adding external data sources 237 such as data obtained from literature is essential for providing a reliable and more accurate tool compared to current prediction tools. This is since data extracted from literature enables to enlarge the amount of data that is relevant to the population of subjects, and thus to
10 serve as big data, which may provide the right impact and influence of each parameter on the final prediction. In addition, some of the data extracted from literature may be universal data, instead of merely geographical data, which may thus assist in applying the data and thus the prediction to different populations at different locations around the world.

[0056] In some embodiments, the at least one meta-analysis parameter 238 may be generated
15 by retrieving data analysis and data classification performed in published articles. For example, the article “Smart Device Use and Perceived Physical and Psychosocial Outcomes among Hong Kong Adolescents 2017” collects data regarding frequency of handphone use in hours per day. The classification used in this article includes 1 to denote no use at all, 2 to denote 1-60 minutes of use per day, 3 to denote 61 to 120 minutes of use per day, 4 to denote 121-180
20 minutes per day, 5 to denote 181-240 minutes per day, and 6 to denote more than 240 minutes per day. The data collected by this article (and others) may be implemented as part of model 240 and may be calculated as part of predicting the risk of developing myopia by model 240. Other articles, that include different data analysis and different data classification may be implemented as part of model 240.

25 [0057] According to some embodiments, at least one database 230 may be used to train the at least one model 240 to provide an indication of a risk of developing myopia. In some embodiments, the at least one model, e.g., at least one model 240 is trained based on individual data arrays generated based on the population data. In some embodiments, the original raw data
30 202 is prepared for modeling, e.g., raw data 202 is used to create feature 210, which are then manipulated into population data arrays 212, and further manipulated into variables 216, while subject data 236 is manipulated into variables 232. The variables are then filtered to those variables that are to be used for training the at least one model 240. The filtered variables may

be split into variables for training the at least one model 240 (e.g., 70% of the variables may be used for training the model), and variables for testing the at least one model 240 (e.g., 30% of the variables may be used for testing the model). The training dataset or training variables may be used to create the at least one model 240 by: (1) converting the data to survival format using survival function; and (2) using a ranger model to train the survival data. The ranger model may produce, e.g., 500 random forest trees and may go through several iterations to improve accuracy of the model. Finally, the model, e.g., model 240, may be tested against the remaining dataset or variables to evaluate the predictive power of the at least one model 240. In some embodiments, metrics such as Rsquared and RMSE may be used for model evaluation. Following training and testing, model 240 may be applied, by the one or more processors (e.g., computer processors 102, Fig. 1), on the received subject data 236, the retrieved at least one population data entry and the generated at least one variable, whereby the at least one variable may comprise variables based on population data arrays and/or subject data.

[0058] Example inputs for at least one model for predicting risk of a subject developing myopia is illustrated in Fig. 3. In some embodiments, the inputs 302 to model 240, may include a subject's current age 312, the subject's gender 314, the subject's eye-test compliance 316, screen time 318, i.e., time subject spends in front of a computerized screen, outdoor time 320, i.e., time the subject spends outdoors, and the subject's parental myopia status 322. The output of model 240 may be a real-time prediction 330, i.e., a future prediction based on real-time inputs, of a risk of the subject to develop myopia. Some of the inputs are inputted at the beginning of the training process as part of the survival function, e.g., current age and eye-test compliance, whereas some inputs are implemented at a later stage as part of the ranger forest model, and some are included as risk factors, after the ranger forest model is complete, e.g., parental myopia status, outdoor time, gender and screen time. The inputs 302 are manipulated to generate variables, and those variables are inputted into model 240 to produce the real-time prediction 330 of the risk of a subject to develop myopia at a current age.

[0059] In some embodiments, the data stored in at least one database 230 may be used to train the at least one model 240. In some embodiments, one or more variables 216 may be selected for such training, based on significance, using one or more algorithms, e.g. Cox proportional-hazards model. Fig. 4A is a schematic illustration of a Ranger survival model plot illustrating a relationship between survival probability of myopia incidence over time, in accordance with embodiments of the present disclosure. While other prediction models make predictions of

whether an event will occur, survival analysis models predict whether the event will occur at a specified time. Thus, a survival model requires a time component for prediction and correspondingly, predicts the time when an event will happen.

[0060] According to Fig. 4A, the survival probability, i.e., the risk of developing myopia is reduced as more time passes. That is, a subject is less likely to develop myopia, as the subject gets older, assuming he is currently non-myopic.

[0061] According to some embodiments, the prediction model 240 (Fig. 2) may implement a Random Survival Forest Ranger Model to train the model. Random Survival Forest Ranger model is an ensemble of tree-based learners that can be used for building a risk prediction model in survival analysis. In survival settings, the predictor may be an ensemble formed by combining results of many survival trees. A Random Survival Forest Ranger model ensures that individual survival trees are de-correlated by: 1) building each tree on a different bootstrap sample of the original training data, and 2) at each node, only evaluating the split criterion for a randomly selected subset of features and thresholds. Predictions are formed by aggregating predictions of individual trees in the ensemble. Ranger is a fast implementation of random forests (Breiman 2001) or recursive partitioning, particularly suited for high dimensional data. Classification, regression, and survival forests are supported by Random Survival Forest Ranger Model. It has a number of speed advantages, including the ability to grow trees in parallel. In some embodiments, the at least one model 240 may implement a Random Survival Forest Ranger Model is fed with certain parameters and variables, which have shown to have a significant effect on the prediction risk value, an example of which is listed hereinbelow. The significance of the certain variables may be determined by feeding the generated variables into a Cox proportional-hazards model. In other embodiments, other variables may be inputted to model 240.

The variables that showed to have a significant effect on the prediction risk value may include: CurrentAge, which denotes a value equal to the current age of a subject for which risk of developing myopia is sought;

MyoGroup, which denotes the binary representation identifying whether the subject has myopia in both eyes;

Screen_time_hr, which denotes a value based on a mean number of hours per day spent by the subject viewing electronic screens or displays;

- Gender, which denotes the gender of the subject;
- EyeTestComp, which denotes a binary representation representing whether the subject has taken an eye test within the past year;
- breakSessAgeTab, which denotes mean number of break sessions from viewing electronic
5 screens or displays, per day;
- breakFreqAgeTab, which denotes mean break frequency per day;
- MaxDiff2, which denotes a derived variable based on an age associated with the highest proportion of population with myopia onset titled OnsetAgeTab, minus the subject eye test age;
- 10 dTree1, which denotes whether MaxDiff2 is more than or equal to 0;
- dTree4, which denotes whether MaxDiff2 is more than or equal to 6;
- dTree5, which denotes whether OnsetAgeTab (which is the proportion of population with myopia onset by current age) is more than or equal to 0.012;
- dTree7, which denotes whether OnsetAgeTab is less than 0.075;
- 15 dTree9, which denotes whether OnsetAgeTab is less than 0.098;
- dTree12, which denotes whether OnsetAgeTab is less than 0.045;
- dTree15, which denotes whether EyeTestDiffTab (which is the mean eye test age difference based on a differential between the current age of each subject and the eye test age of each subject) is more than or equal to 6.7;
- 20 dTree17, which denotes whether EyeTestDiff (assigned a value equal to a difference in years between CurrentAge and EyeTestAge) is more than 8;
- article1, which denotes the article "Early Childhood Refractive Error and Parental History of Myopia as Predictors of Myopia" by Jordan et al 2010, Investigative Ophthalmology & Visual Science, January 2010, Vol. 51, No. 1;

Inter1, which denotes a variable that is based on EyeTestTab (which is the mean eye test age) and MaxDiff2; and

Inter5, which denotes a variable that is based on EyeTestDiffTab and a differential between the subject current age and the subject eye test age.

5 [0062] Fig. 4B is a schematic illustration of a Random Forest Ranger model plot illustrating a relationship between the average survival probability of myopia incidence over time, per all of the parameters implemented in the Random Forest Ranger model, in accordance with embodiments of the present disclosure.

[0063] According to some embodiments, the final prediction model, e.g., prediction model 240
10 (Fig. 2) may further implement a Single Event Probability Calculator using the following formulas:

(ii) $P(E) = n(E) / n(T) = (\text{number of outcomes in the event}) / (\text{total number of possible outcomes})$

(iii) $P(E') = P(\text{not } E) = 1 - P(E)$

whereby

15 $P(E)$ denotes the probability of an Event, in this case the event is developing myopia.

[0064] According to some embodiments, in the final model, the effect of the meta-analysis Relative Risk (RR) output that were previously created, e.g., for creating at least one meta-analysis parameter 238, will be multiplied by risk factors that have not yet been included in the ranger forest model. For example, risk factors that were not included by the Forest Ranger are
20 outdoor activity time, parental myopia, screen time, and eye check compliance. In order to incorporate these risks into the final prediction, ORs (Odds Ratio) were used from literature and meta-analysis techniques were implemented to create a final OR for each of the risk factors that are desired to be used to affect the results of the model. The results of the meta-analysis techniques are also probabilities, thus, these probabilities are simply multiplied by the
25 probability result of the Forest Ranger model.

The RR is the probability that a member of an exposed group will develop a disease relative to the probability that a member of an unexposed group will develop that same disease:

$$(iv) RR = \frac{\text{Probability of getting disease in exposed group}}{\text{Probability of getting disease in non-exposed group}}$$

[0065] Accordingly, the final probability of remaining Non-myopic may be generated from the following formula:

(v) (Parental Myopia meta-analysis RR) × (outdoor Activity RR) × (screen time meta analysis RR) × (eye check compliance RR) × Ranger (probability of remaining non-myopic)

5 In some embodiments, calculation of the final risk of developing myopia will be the survival probability of the Current age of the user.

[0066] In the present disclosure, the rate of survival probability reported in literature was used as a benchmark to determine the accuracy of the disclosed survival model and to further determine if varying the effects of screen time, outdoor activity and parental myopia status,
10 as well as the combination effect of these inputs were consistent with what is reported in literature.

[0067] For example, according to Article 1 ("Early Childhood Refractive Error and Parental History of Myopia as Predictors of Myopia" by Jordan et al 2010, Investigative Ophthalmology & Visual Science, January 2010, Vol. 51, No. 1) high risk of myopia among non-myopic
15 children was defined as hyperopia of 0.75 D or less hyperopia in the cycloplegic sphere (in negative cylinder convention) in the first grade. Article 1 teaches that assuming a child has more than one parent with myopia, as a child grows older, the probability of remaining non-myopic decreases, whereas if the child's parents are non-myopic, as the child grows older, the probability of remaining non-myopic decreases but at a lower rate than that of a child
20 with myopic parents.

[0068] Fig. 5 is a schematic illustration of a method 500 for predicting a risk of developing myopia, in accordance with embodiments of the present disclosure. Method 500 may be carried out by prediction module 122 (Fig. 1) or prediction module 200 (Fig. 2) and/or data analysis module 126 (Fig. 1). In some embodiments, method 500 may begin with operation 502 wherein
25 computer system 100 comprising prediction module 200 receives subject data 236 on a subject via one or more computer processors, e.g., computer processor 102. The computer system may comprise one or more computer processors, and a memory storing at least one program for execution by the one or more processors.

[0069] Receiving subject data 502 may comprise receiving a subject current age, a subject
30 gender, a subject eye test compliance data, a subject screen time data comprising frequency of exposure of the subject to electronic screens, a subject outdoor time data comprising time spent

by the subject outdoors, a subject parental myopia status (i.e., that either none of the subject's parents are myopic, one of the parents is myopic or both parents are myopic).

[0070] Method 500 may further comprise operation 504, wherein at least one population data array, e.g., population data array 212 (Fig. 2) generated based on a population data, e.g., population data 226, and stored in at least one database, e.g., database 230, is accessed by the one or more computer processors. In some embodiments, the population data comprises data from different subjects and corresponding instances of myopia that occurred to the different subjects. Method 500 may further comprise operation 506, wherein at least one population data array entry is retrieved from the at least one population data array.

[0071] Method 500 may comprise operation 508, wherein at least one variable, which is based on at least the received subject data is generated. The generated variable, e.g., generated variable 232, may be based on the received subject data, and may also be based on information stored in at least one database 230, which comprises information related to population data arrays 212 and variable derived therefrom, e.g., derived variables 216.

[0072] Method 500 may comprise operation 510, wherein the risk of a subject to develop myopia is determined by applying one or more trained models based on individual data arrays (e.g., initial individual data arrays 214) generated based on the population data (e.g., population data 226). The one or more trained models may be applied to the received subject data, the retrieved at least one population data array entry and the generated at least one variable. The one or more trained models, e.g., prediction model 240, may be configured to provide an indication of a risk of the subject developing myopia or the probability that the subject will develop myopia. Method 500 may optionally comprise operation 512, wherein at least one risk ratio is applied as part of operation 510. The risk ratio may be a probability of risk factors that were not included by the trained model. For example, in some embodiments, risk factors such as outdoor activity time, parental myopia, screen time, and eye check compliance may be applied according to formula (v) hereinabove.

[0073] Method 500 may comprise operation 514, wherein a score indicating likelihood of the subject to develop myopia is generated. The score may be generated based on the determined risk of the subject to develop myopia. The score may be represented by a number, e.g., percentage, or by some sort of gauge or bar indicating low, medium or high level of likelihood

of the subject to develop myopia at a current age. Other score representations may be implemented.

[0074] Method 500 may comprise operation 516, wherein a report comprising the score indicating likelihood of the subject to develop myopia is generated. In some embodiments, the report may comprise an actual recommendation for actions that may be performed to decrease the likelihood or risk of developing myopia.

[0075] In some embodiments, method 500 may comprise modifying the score, by the one or more computer processors, based on updated received subject data and updated received population data array over time. That is, the score or risk of the subject to develop myopia may be continuously updated, in real-time, as more updated subject related data is collected and received by the one or more computer processors, and as more updated population related data is received by the one or more computer processors.

[0076] In some embodiments, method 500 may comprise storing the risk of the subject to develop myopia. The stored risk may be used for future calculations and training of the model.

[0077] According to some embodiments, the at least one variable may be generated by generating at least one individual data array comprising at least one data point for each different subject; generating a plurality of decision trees based on the at least one individual data array and identifying a plurality of decision tree nodes; identifying at least one interaction between the plurality of decision trees; and selecting at least one variable from the at least one data point, the plurality of decision tree nodes, and the at least one interaction, based on significance.

[0078] According to some embodiments, the significance of a variable and input parameter, such as the subject data 236, and/or population data 226, may be established by calculating the risk of developing myopia, while changing a certain variable or input parameter, and determining whether the change of that variable or parameter had any effect of the risk value.

[0079] In some embodiments, the following tables illustrate that the trained model, e.g., trained model 240, works, by showing that the parameters affect the outcome of the risk prediction. Table 6 illustrates the risk calculated with a changing age, while all the other parameters are with a minimum value.

Table 6: Effect of age on Risk of developing myopia

Gender	Eye Test	Age	Outdoor	Parent	Screen time	RISK
Male	No	4	0	None	0	0.61
Male	No	6	0	None	0	0.54
Male	No	8	0	None	0	0.46
Male	No	10	0	None	0	0.37
Male	No	12	0	None	0	0.33

Table 6 teaches that the risk of developing myopia decreases with age.

5 [0080] Table 7 illustrates the effect of outdoor activity (hours per week) on the risk model.

Table 7: Effect of time of outdoor activity on Risk of developing myopia

Gender	Eye Test	Age	Outdoor	Parent	Screen time	RISK
Male	No	4	5	None	0	0.47
Male	No	6	5	None	0	0.42
Male	No	8	5	None	0	0.36
Male	No	10	5	None	0	0.29
Male	No	12	5	None	0	0.25
Gender	Eye Test	Age	Outdoor	Parent	Screen time	RISK
Male	No	4	15	None	0	0.31
Male	No	6	15	None	0	0.28
Male	No	8	15	None	0	0.24
Male	No	10	15	None	0	0.19
Male	No	12	15	None	0	0.17

Table 7 teaches that the calculated model risk decreases with increasing outdoor time.

[0081] Table 8 illustrates the effect of screen time (hours per day) on the risk model.

Table 8: Effect of screen time on Risk of developing myopia

Gender	Eye Test	Age	Outdoor	Parent	Screen time	RISK
Male	No	4	0	None	2	0.67
Male	No	6	0	None	2	0.59
Male	No	8	0	None	2	0.51
Male	No	10	0	None	2	0.47
Male	No	12	0	None	2	0.41
Gender	Eye Test	Age	Outdoor	Parent	Screen time	RISK
Male	No	4	0	None	6	0.95
Male	No	6	0	None	6	0.9
Male	No	8	0	None	6	0.78
Male	No	10	0	None	6	0.61
Male	No	12	0	None	6	0.55

Table 8 teaches that the calculated model risk increases with increasing screen time.

5 [0082] Similar calculations may be done for determining effect of parental myopia status on the risk model (in such case, the calculated model risk increases with greater incidence of parental myopia), determining effect of eye test on risk model (in which case the calculated model risk decreases very slightly with eye test compliance), effect of gender on the risk model . In some embodiments, gender related literature illustrates a slight effect such that females are slightly more of risk to develop myopia, compared to males. This gender related literature may be added as part of the risk factors in equation (v).

10

[0083] In some embodiments, the at least one population data array may comprise a first population data array, the first population data array may comprise at least one age and a mean break session data associated with each of the at least one age, the break session data relating to a duration of break sessions from viewing electronic screens or displays taken by each subject per day. In some embodiments, the at least one population data array entry may comprise a first population data array entry [e.g., breakSessAgeTab], the first population data array entry may comprise the mean break session data associated with the at least one age corresponding to the subject current age.

15

[0084] In some embodiments, the at least one population data array may comprise a second population data array, which may comprise at least one age and a mean break frequency data associated with each of the at least one age, the break frequency data relating to a frequency of break sessions from viewing electronic screens or displays taken by each subject per day. In
5 some embodiments, the at least one population data array entry may comprise a second population data array entry [e.g., breakFreqAgeTab], the second population data array entry may comprise the mean break frequency data associated with the at least one age corresponding to the subject current age.

[0085] In some embodiments, the at least one generated variable may comprise a first variable,
10 which may comprise a subject eye test age based on the subject current age and the subject eye test compliance data.

[0086] In some embodiments, the at least one population data array may comprise a third population data array, the third population data array may comprise at least one age and a proportion of population with myopia onset associated with each of the at least one age [e.g.,
15 OnsetAgeTab].

[0087] In some embodiments, the at least one population data array entry may comprise a third population data array entry [e.g., max onset age], the third population data array entry comprising the at least one age associated with the highest proportion of population with myopia onset.

20 [0088] In some embodiments, the at least one variable may comprise a second variable [e.g., MaxDiff2], the second variable may be generated based on the third population data array entry [e.g., max onset age] and the subject eye test age [eyetestage].

[0089] In some embodiments, the at least one variable may be selected from the group comprising of: a third variable [dtree1], the third variable based on whether the second variable
25 is more than or equal to 0; and

a fourth variable [dtree4], the fourth variable based on whether the second variable is more than or equal to 6.

[0090] In some embodiments, the at least one population data array may further comprise a fourth population data array, the fourth population data array may comprise at least one age
30 and a mean eye test age associated with each of the at least one age. In some embodiments, the

at least one population data array entry may further comprise a fourth population data array entry [e.g., EyeTestTab], the fourth population data array entry comprising a mean eye test age associated with the at least one age corresponding to the subject current age.

5 [0091] In some embodiments, the at least one variable may comprise a fifth variable [e.g., inter1] based on the fourth population data array entry [e.g., EyeTestTab] and the second variable [e.g., MaxDiff2].

[0092] In some embodiments, the at least one population data array entry may comprise a fifth population data array entry, the fifth population data array entry may comprise a proportion of population with myopia onset associated with the at least one age corresponding to the subject
10 current age [e.g., proportion from OnsetAgeTab].

[0093] In some embodiments, the at least one variable may be selected from the group comprising of:

a sixth variable based on whether the fifth variable is more than or equal to 0.012 [e.g., dtree5];
a seventh variable based on whether the fifth variable is less than 0.075 [e.g., dtree7];
15 an eighth variable based on whether the fifth variable is less than 0.098 [e.g., dtree9];
a ninth variable based on whether fifth variable is less than 0.045 [e.g., dtree12].

[0094] In some embodiments, the at least one variable may comprise a tenth variable [e.g., dtree17], the tenth variable may be based on whether a differential [EyeTestDiff] between the current age and the first variable [e.g., eyetestage] is more than 8.

20 [0095] In some embodiments, the at least one population data array may comprise a fifth population data array [e.g., EyeTestDiffTab], the fifth population data array may comprise at least one age and a mean eye test age difference associated with each of the at least one age, the mean eye test age difference is based on a mean difference between an age and an eye test age of the population data [e.g., eyetestdiff].

25 [0096] In some embodiments, the at least one variable may comprise an eleventh variable [e.g., dtree15], the eleventh variable may be based on whether the fifth population data array entry is more than or equal to 6.7.

- [0097] In some embodiments, the at least one variable may comprise a twelfth variable [e.g., inter5], the twelfth variable may be based on the fifth population data array entry and a differential between the subject current age and the subject eye test age.
- 5 [0098] In some embodiments, the at least one variable may comprise a thirteenth variable, the thirteenth variable may be based on the subject screen time data [e.g., data extracted from article1].
- [0099] In some embodiments, method 500 may further comprise applying at least one risk ratio based on the received parental myopia status data, e.g., as in operation 512.
- 10 [0100] In some embodiments, method 500 may comprise applying at least one risk ratio based on received outdoor data, e.g., as in operation 512.
- [0101] In some embodiments, method 500 may comprise applying at least one risk ratio based on the subject gender data, as in operation 512.
- [0102] In some embodiments, method 500 may comprise applying at least one risk ratio based on the subject screen time data, i.e., based on the amount of time, e.g., hours per day that the
15 subject is spending watching electronic screens or displays, as in operation 512.
- [0103] In some embodiments, method 500 may comprise applying at least one risk ratio based on the subject eye test compliance data, i.e., based on the number, time, frequency of eye check-ups the subject underwent and/or the age of the subject when undergoing such eye check-ups, as in operation 512.
- 20 [0104] In some embodiments, method 500 may comprise an on-going continuous training of the at least one model 240, by a cumulative collection of population data and subject data. Such continuous cumulative growth of the dataset strengthens the model, since the model is re-trained every period of time based on a larger dataset, which in turn provides a more accurate prediction.
- 25 [0105] In some embodiments, the prediction module, e.g. prediction module 122 or prediction module 200, may receive input from several sources: (a) a mobile application installed on a computerized user device, e.g., a mobile phone, a tablet, a wearable, a computer, or any other computerized device, for example, plano®App, by Plano Pte Ltd.; (b) a standalone designated PMCV (Predict My Child Vision) user-facing platform, i.e., a user-facing platform designed

specifically for collecting data for prediction module 122 or 200, of the present disclosure; (c) an eye-check platform, e.g., the planoEyecheck platform; and (d) scientific literature. In some embodiments, some of the abovementioned input platforms may also serve as output platforms, which may, for example, display the report comprising the calculated risk or score indicating
5 likelihood of a subject to develop myopia, as well as provide prescriptive aspects, such as recommendations for the subject that can assist in reducing the risk of developing myopia.

[0106] Conjugated terms such as, by way of example, 'a thing property' implies a property of the thing, unless otherwise clearly evident from the context thereof.

[0107] The flowchart and block diagrams illustrate architecture, functionality or an operation
10 of possible implementations of systems, and methods according to various embodiments of the present disclosed subject matter. It should also be noted that, in some alternative implementations, illustrated or described operations may occur in a different order or in combination or as concurrent operations instead of sequential operations to achieve the same or equivalent effect.

[0108] The corresponding structures, materials, acts, and equivalents of all means or step plus
15 function elements in the claims below are intended to include any structure, material, or act for performing the function in combination with other claimed elements as specifically claimed. As used herein, the singular forms "a", "an" and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise. It will be further understood that the
20 terms "comprising", "including" and/or "having" and other conjugations of these terms, when used in this specification, specify the presence of stated features, integers, steps, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features, integers, steps, operations, elements, components, and/or groups thereof.

[0109] The terminology used herein should not be understood as limiting, unless otherwise
25 specified, and is for the purpose of describing particular embodiments only and is not intended to be limiting of the disclosed subject matter. While certain embodiments of the disclosed subject matter have been illustrated and described, it will be clear that the disclosure is not limited to the embodiments described herein. Numerous modifications, changes, variations, substitutions and equivalents are not precluded.

CLAIMS

1. A computer-implemented method for predicting a risk of developing myopia in a subject, comprising:
 - at a computer system having one or more computer processors, and a memory storing
 - 5 at least one program for execution by the one or more computer processors:
 - receiving, by the one or more computer processors via a computerized device,
 - a subject data on a subject, wherein the subject data comprises:
 - a subject current age;
 - a subject gender;
 - 10 a subject country of residence;
 - a subject eye test compliance data;
 - a subject screen time data;
 - a subject outdoor data;
 - a subject parental myopia status;
 - 15 accessing, by the one or more computer processors, at least one population data array generated based on a population data, the population data comprising data received from different subjects and corresponding instances of myopia that occurred for the different subjects, said population data array is stored in at least one database;
 - retrieving by the one or more computer processors, at least one population data
 - 20 array entry from the at least one population data array;
 - generating, by the one or more computer processors, at least one variable based on at least the received subject data;
 - determining the risk of the subject to develop myopia, by applying at least one model, by the one or more processors, on the received subject data, the at least one
 - 25 population data entry and the generated at least one variable, said model is trained based on individual data arrays generated based on the population data;
 - generating a score indicating a likelihood of the subject to develop myopia, based on the determined risk; and
 - generating a report comprising the score indicating the likelihood of the subject
 - 30 to develop myopia.

2. The computer-implemented method of claim 1, further comprising modifying the score, by the one or more computer processors, based on updated received subject data and updated received population data array over time.
3. The computer-implemented method of claim 1, further comprising storing the risk of the subject to develop myopia.
4. The computer-implemented method of claim 1, wherein generating the at least one variable comprises:
 - generating at least one individual data array comprising at least one data point for each different subject;
 - generating a plurality of decision trees based on the at least one individual data array and identifying a plurality of decision tree nodes;
 - identifying at least one interaction between the plurality of decision trees; and
 - selecting at least one variable from the at least one data point, the plurality of decision tree nodes, and the at least one interaction, based on significance.
5. The computer-implemented method of claim 1, wherein:
 - the at least one population data array comprises a first population data array, the first population data array comprising at least one age and a mean break session data associated with each of the at least one age; and
 - the at least one population data array entry comprises a first population data array entry, the first population data array entry comprising the mean break session data associated with the at least one age corresponding to the subject current age.
6. The computer-implemented method of claim 1, wherein:
 - the at least one population data array comprises a second population data array, the second population data array comprising at least one age and a mean break frequency data associated with each of the at least one age; and
 - the at least one population data array entry comprises a second population data array entry, the second population data array entry comprising the mean break frequency data associated with the at least one age corresponding to the subject current age.

7. The computer-implemented method of claim 1, wherein the at least one variable comprises a first variable, the first variable comprising a subject eye test age based on the subject current age and the subject eye test compliance data.
8. The computer-implemented method of claim 1, wherein the at least one population data
5 array comprises a third population data array, the third population data array comprising at least one age and a proportion of population with myopia onset associated with each of the at least one age.
9. The computer-implemented method of claim 1, wherein the at least one population data array entry comprises a third population data array entry, the third population data array entry
10 comprising an at least one age associated with the highest proportion of population with myopia onset.
10. The computer-implemented method of claim 9, wherein the at least one variable comprises a second variable based on a chronological age at which myopia first appeared, the second variable generated based on the third population data array entry and a subject eye test
15 age.
11. The computer-implemented method of claim 10, wherein the at least one variable is selected from the group comprising of:
- a third variable, the third variable based on whether the second variable is more than or is equal to 0; and
- 20 a fourth variable, the fourth variable based on whether the second variable is more than or is equal to 6.
12. The computer-implemented method of claim 10, wherein:
- the at least one population data array further comprises a fourth population data array, the fourth population data array comprising at least one age and a mean eye test age associated
25 with each of the at least one age; and
- the at least one population data array entry further comprises a fifth population data array entry, the fifth population data array entry comprising a mean eye test age associated with the at least one age corresponding to the subject current age.
13. The computer-implemented method of claim 12, wherein the at least one variable
30 comprises a fifth variable based on a fourth population data array entry and the second variable.

14. The computer-implemented method of claim 1, wherein the at least one population data array entry comprises a fifth population data array entry, the fifth population data array entry comprising a proportion of population with myopia onset associated with an at least one age corresponding to the subject current age.
- 5 15. The computer-implemented method of claim 14, wherein the at least one variable is selected from the group comprising of:
- a sixth variable based on whether a fifth variable is more than or equal to 0.012;
 - a seventh variable based on whether the fifth variable is less than 0.075;
 - an eighth variable based on whether the fifth variable is less than 0.098; and
 - 10 a ninth variable based on whether the fifth variable is less than 0.045.
16. The computer-implemented method of claim 1, wherein the at least one variable comprises a tenth variable, the tenth variable based on whether a differential between the current age and a first variable is more than 8.
17. The computer-implemented method of claim 1, wherein:
- 15 the at least one population data array comprises a fifth population data array, the fifth population data array comprising at least one age and a mean eye test age difference associated with each of the at least one age, the mean eye test age difference based on a differential between a user current age and an eye test age of the population data; and
- the at least one variable comprises an eleventh variable, the eleventh variable based on
- 20 whether a fifth population data array entry is more than or equal to 6.7.
18. The computer-implemented method of claim 17, wherein the at least one variable comprises a twelfth variable, the twelfth variable based on a fifth population data array entry and a differential between the subject current age and a subject eye test age.
19. The computer-implemented method of claim 1, wherein the at least one variable
- 25 comprises a thirteenth variable, the thirteenth variable based on the subject screen time data.
20. The computer-implemented method of claim 1, further comprising applying at least one risk ratio based on the subject parental myopia status.
21. The computer-implemented method of claim 1, further comprising applying at least one risk ratio based on the subject outdoor data.

22. The computer-implemented method of claim 1, further comprising applying at least one risk ratio based on the subject gender.
23. The computer-implemented method of claim 1, further comprising applying at least one risk ratio based on the subject screen time data.
- 5 24. The computer-implemented method of claim 1, further comprising applying at least one risk ratio based on the subject eye test compliance data.
25. The computer-implemented method of claim 1, wherein the generated report comprises recommendations to a user on actions to reduce risk of the subject developing myopia.