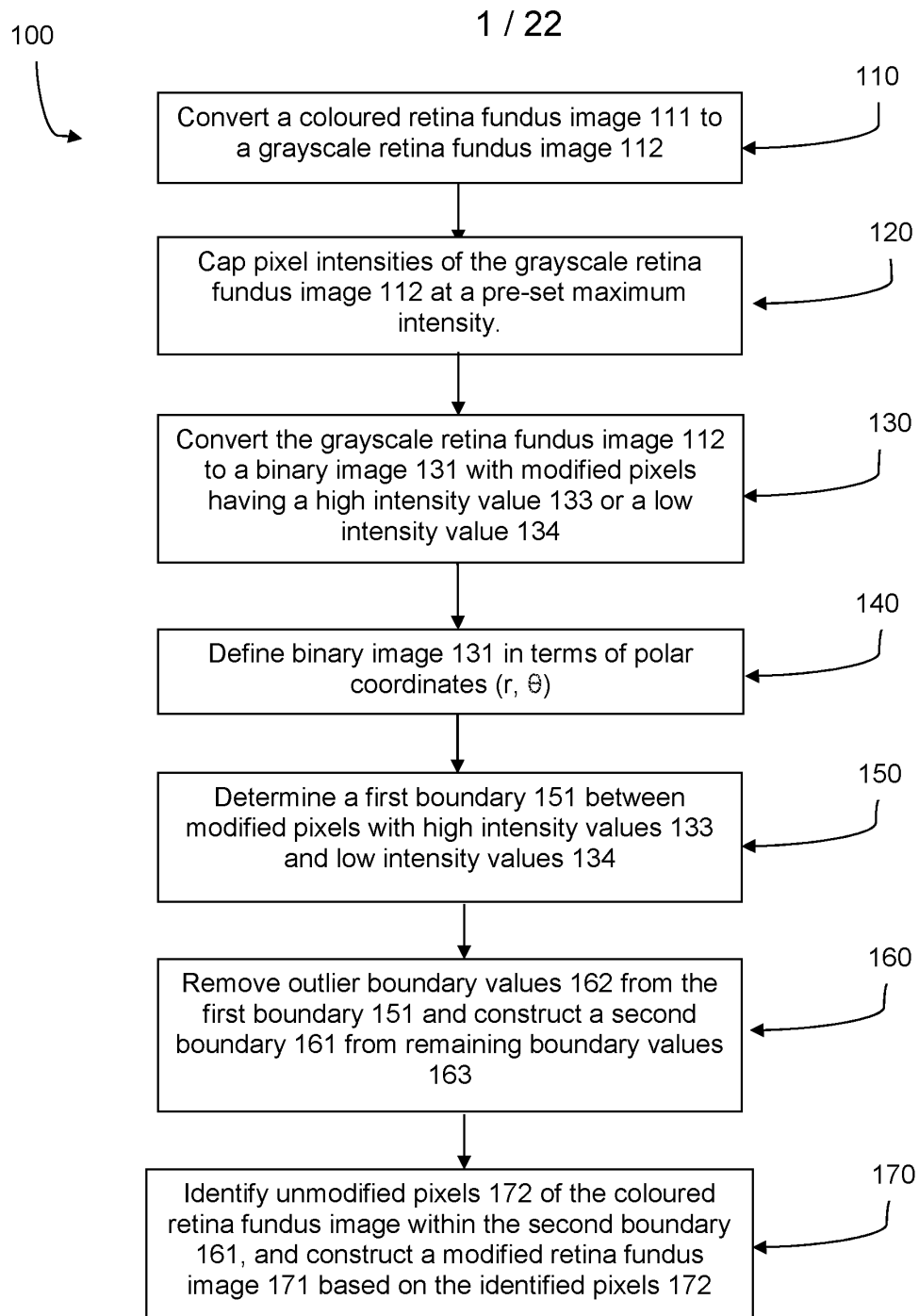
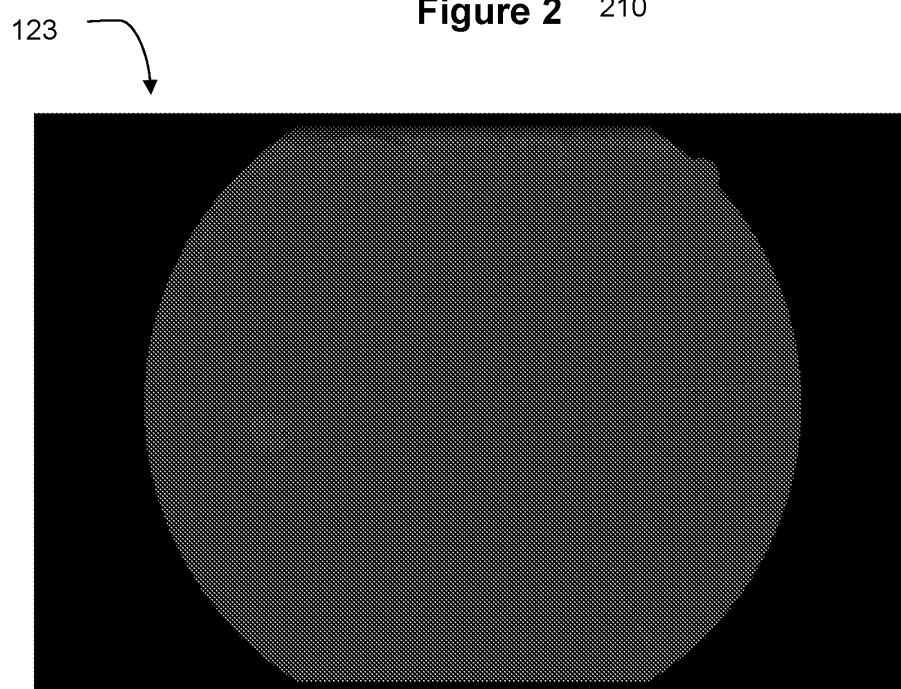
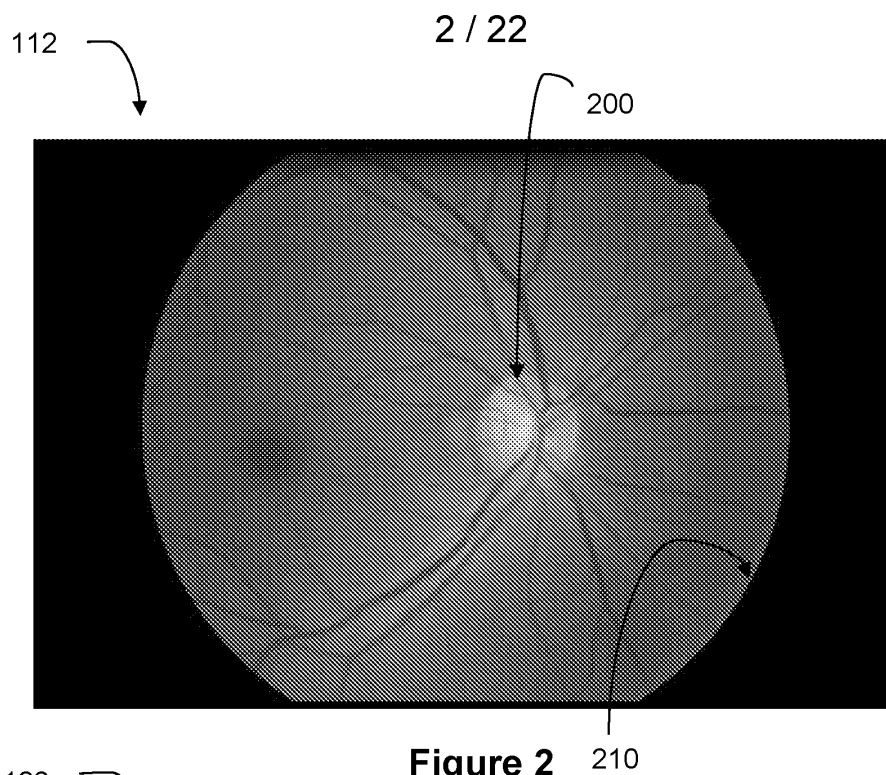


**Figure 1**



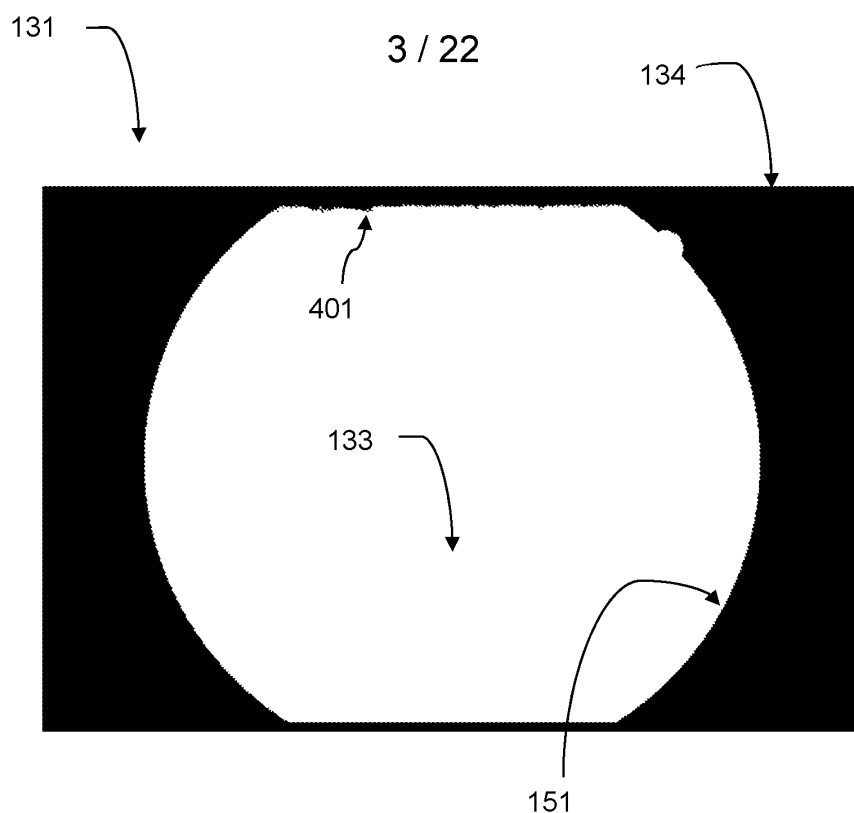


Figure 4

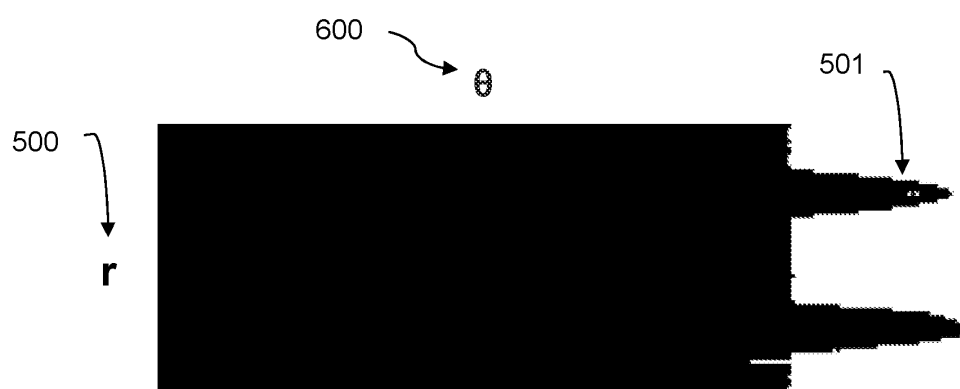


Figure 5

4 / 22

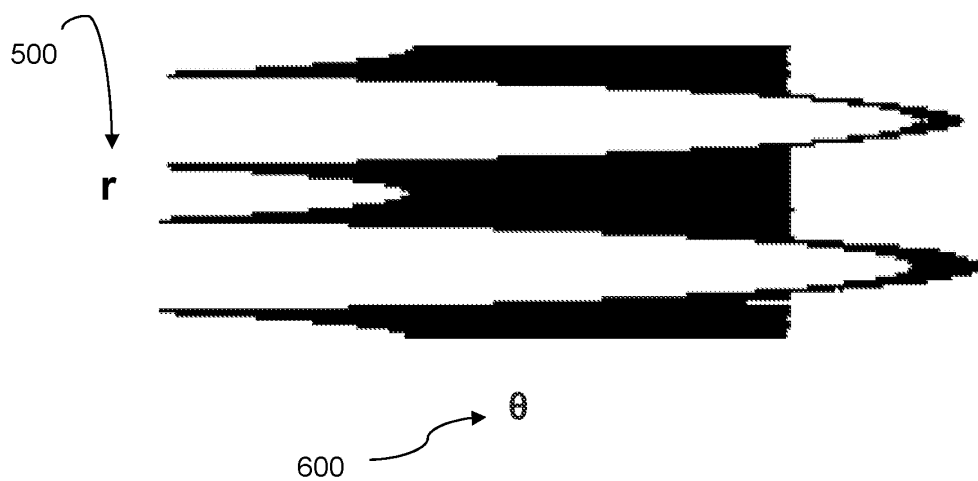


Figure 6

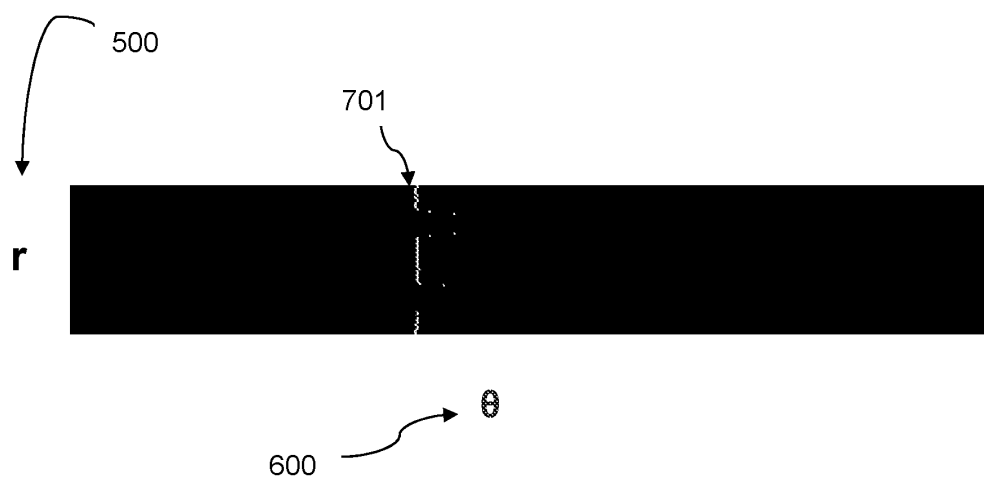


Figure 7

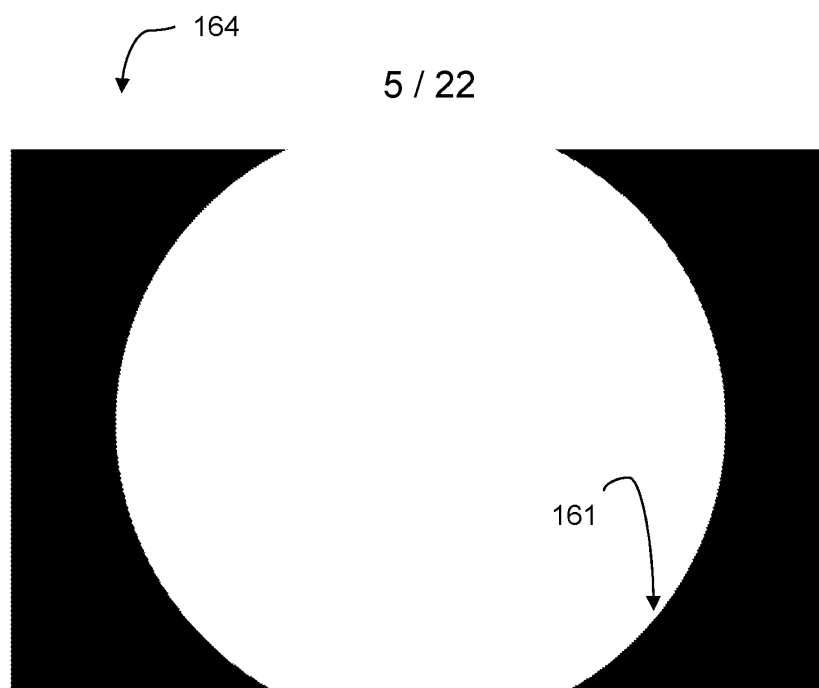


Figure 8

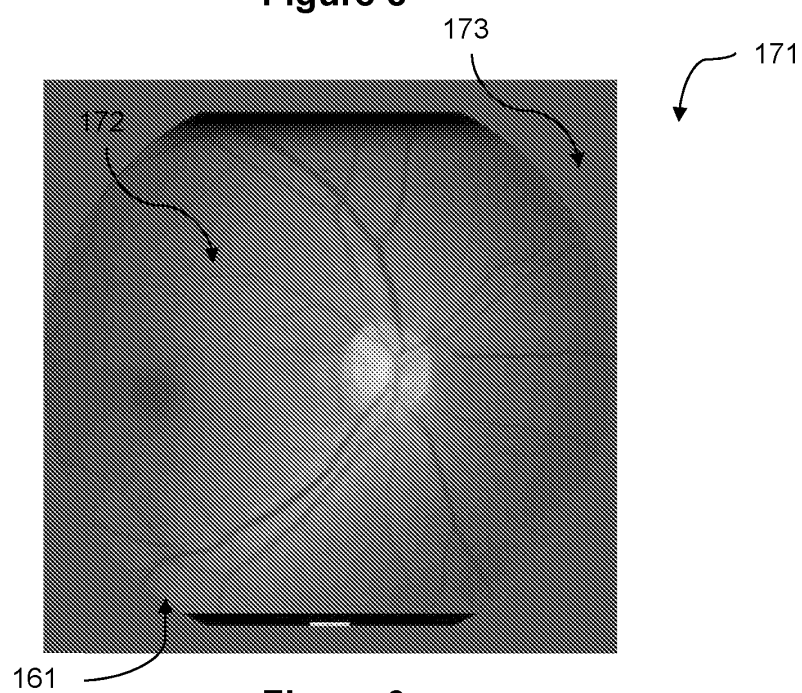


Figure 9

6 / 22

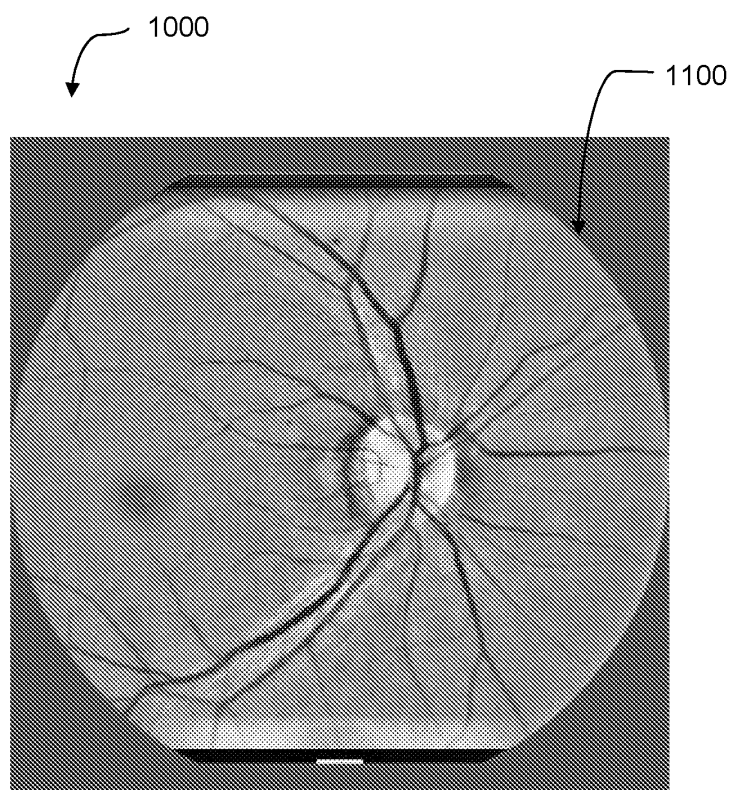


Figure 10

**Table 1: Summary of the training and validation datasets for referable diabetic retinopathy (DR), glaucoma suspect (GS) and age-related macular degeneration (AMD)**

| <b>Referable Conditions</b>         | <b>Dataset</b>              | <b>Countries</b> | <b>Types of cohorts</b> | <b>Ethnicities</b>     | <b>Images</b> |
|-------------------------------------|-----------------------------|------------------|-------------------------|------------------------|---------------|
| <i>Referable DR</i>                 |                             |                  |                         |                        |               |
| <b>Training dataset</b>             | SIDRP 2010-13               | Singapore        | Community-based         | Chinese, Malay, Indian | 76,370        |
| <b>Primary validation dataset</b>   | SIDRP 2014-15               | Singapore        | Community-based         | Chinese, Malay, Indian | 71,896        |
| <b>External validation datasets</b> | Zhongshan Ophthalmic Center | China            | Community-based         | Chinese                | 15,798        |
|                                     | SIMES                       | Singapore        | Population-based        | Malay                  | 1,936         |
|                                     | SINDI                       | Singapore        | Population-based        | Indian                 | 3,052         |
|                                     | SCES                        | Singapore        | Population-based        | Chinese                | 4,512         |
|                                     | BES                         | China            | Population-based        | Chinese                | 1,052         |
|                                     | AFEDS                       | USA              | Population-based        | African                | 1,968         |
|                                     | DMP <sup>49</sup>           | Australia        | Clinic-based            | Caucasian (white)      | 2,302         |
|                                     | Mexican Eye Hospital        | Mexico           | Clinic-based            | Hispanics              | 1,172         |
|                                     | CUHK                        | Hong Kong        | Clinic-based            | Chinese                | 1,254         |
|                                     | HKU                         | Hong Kong        | Clinic-based            | Chinese                | 7,706         |
|                                     |                             |                  |                         | <b>Total</b>           | 189,018       |
| <i>Referable GS</i>                 |                             |                  |                         |                        |               |
| <b>Training dataset</b>             | SIDRP 2010-13               | Singapore        | Community-based         | Chinese, Malay, Indian | 76,108        |
|                                     | SIMES                       | Singapore        | Population-based        | Malay                  | 10,114        |
|                                     | SINDI                       | Singapore        | Population-based        | Indian                 | 10,819        |
|                                     | SCES                        | Singapore        | Population-based        | Chinese                | 26,731        |
|                                     | SNEC Glaucoma study         | Singapore        | Clinic-based            | Chinese, Malay, Indian | 1,417         |

**Figure 11A**

|                                   |                                                           |                                                                |                                                                                             |                                                                                |                                             |
|-----------------------------------|-----------------------------------------------------------|----------------------------------------------------------------|---------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|---------------------------------------------|
| <b>Primary validation set</b>     | SIDRP 2014-15                                             | Singapore                                                      | Community-based                                                                             | Chinese, Malay, Indian<br><b>Total</b>                                         | 71,896<br>197,085                           |
| <u>Referable AMD</u>              |                                                           |                                                                |                                                                                             |                                                                                |                                             |
| <b>Training dataset</b>           | SIDRP 2010-13<br>SIMES<br>SINDI<br>SCEs<br>SNEC AMD study | Singapore<br>Singapore<br>Singapore<br>Singapore<br>Singapore  | Community-based<br>Population-based<br>Population-based<br>Population-based<br>Clinic-based | Chinese, Malay, Indian<br>Malay<br>Indian<br>Chinese<br>Chinese, Malay, Indian | 36,561<br>8,616<br>7,447<br>16,812<br>2,180 |
| <b>Primary validation dataset</b> | SIDRP 2014-15                                             | Singapore                                                      | Community-based                                                                             | Chinese, Malay, Indian<br><b>Total</b>                                         | 35,948<br>107,564                           |
|                                   |                                                           | <b>Total images for DR, GS and AMD training and validation</b> |                                                                                             |                                                                                | <b>493,667</b>                              |

SIDRP 2010-13: Singapore National Diabetic Retinopathy Screening Program 2010 to 2013  
SIDRP 2014-15: Singapore National Diabetic Retinopathy Screening Program 2014 to 2015  
SIMES: Singapore Malay Eye Study  
SINDI: Singapore Indian Eye Study  
SCEs: Singapore Chinese Eye Study  
SNEC: Singapore National Eye Center  
BES: Beijing Eye Study  
AFEDS: African American Eye Disease Study  
DMP: Diabetes Management Project  
CUHK: Chinese University of Hong Kong  
HKU: University of Hong Kong

Figure 11B

Table 2: Training and validation datasets for diabetic retinopathy (DR)

| Datasets                                                                                                               | Number of patients | Number of images | Number of eyes | No DR (N, %)   | Mild NPDR (N, %) | Moderate NPDR (N, %) | Severe NPDR (N, %) | PDR (N, %)  | DME (N, %)   | Ungradable (N, %) |
|------------------------------------------------------------------------------------------------------------------------|--------------------|------------------|----------------|----------------|------------------|----------------------|--------------------|-------------|--------------|-------------------|
| <b>Training</b><br>SIDRP 2010-13                                                                                       | 13,099             | 76,370           | 38,185         | 33,709 (88.3%) | 3,310 (8.7%)     | 597 (1.6%)           | 478 (1.3%)         | 70 (0.2%)   | 2,026 (5.3%) | 21 (0.1%)         |
| <b>Primary Validation</b><br>SIDRP 2014-15*                                                                            | 14,880             | 71,896           | 35,948         | 33,485 (92.3%) | 1,813 (5.0%)     | 455 (1.3%)           | 171 (0.5%)         | 24 (0.1%)   | 320 (0.9%)   | 0 (0.0%)          |
| <b>External Validation</b><br><i>Community-based</i><br>1. Chinese **<br>(Giangdong, China)<br><i>Population-based</i> | 3,970              | 15,798           | 7,899          | 5,665 (71.7%)  | 1,235 (15.6%)    | 737 (9.3%)           | 0 (0%)             | 154 (1.9%)  | 0 (0%)       | 108 (1.4%)        |
| 2. Chinese (Singapore)                                                                                                 | 484                | 1,936            | 968            | 759 (78.4%)    | 131 (13.5%)      | 60 (6.2%)            | 1 (0.1%)           | 7 (0.7%)    | 17 (1.8%)    | 10 (1%)           |
| 3. Malay (Singapore)                                                                                                   | 763                | 3,052            | 1,526          | 1,143 (74.9%)  | 215 (14.1%)      | 113 (7.4%)           | 18 (1.2%)          | 9 (0.6%)    | 53 (3.5%)    | 28 (1.8%)         |
| 4. Indian (Singapore)                                                                                                  | 1,128              | 4,512            | 2,256          | 1,639 (72.7%)  | 422 (18.7%)      | 125 (5.5%)           | 5 (0.2%)           | 17 (0.8%)   | 71 (3.1%)    | 48 (2.1%)         |
| 5. Chinese (Beijing, China)                                                                                            | 263                | 1,052            | 526            | 493 (93.7%)    | 4 (0.8%)         | 11 (2.1%)            | 4 (0.8%)           | 0 (0%)      | 12 (2.3%)    | 2 (0.4%)          |
| 6. African Americans (AFEDS)                                                                                           | 492                | 1,968            | 984            | 807 (82.0%)    | 50 (5.1%)        | 37 (3.8%)            | 5 (0.5%)           | 16 (1.6%)   | 28 (2.85%)   | 41 (4.17%)        |
| <i>Clinic-based</i><br>7. Caucasian (Australia)                                                                        | 588                | 2,302            | 1,151          | 432 (37.5%)    | 121 (10.5%)      | 159 (13.8%)          | 123 (10.7%)        | 191 (16.6%) | 249 (21.6%)  | 125 (10.9%)       |
| 8. Hispanic (Mexico)                                                                                                   | 343                | 1,172            | 586            | 38 (6.5%)      | 284 (48.5%)      | 192 (32.8%)          | 51 (8.7%)          | 18 (3.1%)   | 223 (38.1%)  | 3 (0.5%)          |

Figure 12A

|                       |        |         |        |                 |                |                |             |              |                |              |
|-----------------------|--------|---------|--------|-----------------|----------------|----------------|-------------|--------------|----------------|--------------|
| 9. Chinese (CUHK)     | 314    | 1,254   | 627    | 224 (35.7%)     | 114 (18.2%)    | 235 (37.5%)    | 43 (6.9%)   | 11 (1.8%)    | 96 (15.3%)     | 0 (0.0%)     |
| 10. Chinese *** (HKU) | 1,932  | 7,706   | 3,853  | 1,984 (51.5%)   | 1,485 (38.5%)  | 155 (4.0%)     | 14 (0.4%)   | 0 (0%)       | 214 (5.5%)     | 1 (0.03%)    |
| Total                 | 38,256 | 189,018 | 94,509 | 80,378 (85.0 %) | 9,184 (9.72 %) | 2,876 (3.04 %) | 913 (0.97%) | 517 (0.55 %) | 3,309 (3.50 %) | 387 (0.41 %) |

\* 6291 patients in SIDRP 2014-2015 were repeats from SIDRP 2010-2013 and 8,589 were unique patients

\*\* 41 patients had only 1 eye

\*\*\* 11 patients had only 1 eye

DR: Diabetic retinopathy; NPDR: Non-proliferative DR; PDR: Proliferative DR; DME: Diabetic macular edema; CUHK: Chinese University of Hong Kong; HKU: The University of Hong Kong

Figure 12B

Table 3: Training set for referable glaucoma suspect (GS) and age-related macular degeneration (AMD)

| Referable Glaucoma Suspect (GS)                                  | Total Images with GS assessment  | No. of image with referable GS   |
|------------------------------------------------------------------|----------------------------------|----------------------------------|
| Singapore Diabetic Retinopathy Screening Program (SIDRP) 2010-13 | 76,108                           | 120                              |
| Singapore Chinese Eye Study                                      | 26,731                           | 603                              |
| Singapore Malay Eye Study                                        | 10,114                           | 333                              |
| Singapore Indian Eye Study                                       | 10,819                           | 157                              |
| Singapore National Eye Center Total GS Images                    | 1417                             | 1,417                            |
|                                                                  | 125,189                          | 2,630                            |
| Referable Age-related Macular Degeneration (AMD)                 | Total Images with AMD assessment | No. of images with referable AMD |
| Singapore Diabetic Retinopathy Screening Program (SIDRP) 2010-13 | 36,561                           | 19                               |
| Singapore National Eye Center                                    | 2,180                            | 1,632                            |
| Singapore Chinese Eye Study                                      | 16,182                           | 43                               |
| Singapore Malay Eye Study                                        | 8,616                            | 26                               |
| Singapore Indian Eye Study                                       | 7,447                            | 18                               |
| Total AMD Images                                                 | 70,986                           | 1,738                            |

GS: Glaucoma suspect; AMD: age-related macular degeneration  
Referable GS – defined as vertical CDR 0.8 and above, local neuro-retinal rim thinning, focal notching, disc haemorrhage, retinal nerve fibre layer defect; Referable AMD – defined as intermediate AMD and/or advanced AMD (geography atrophy and neovascular AMD).

Figure 13

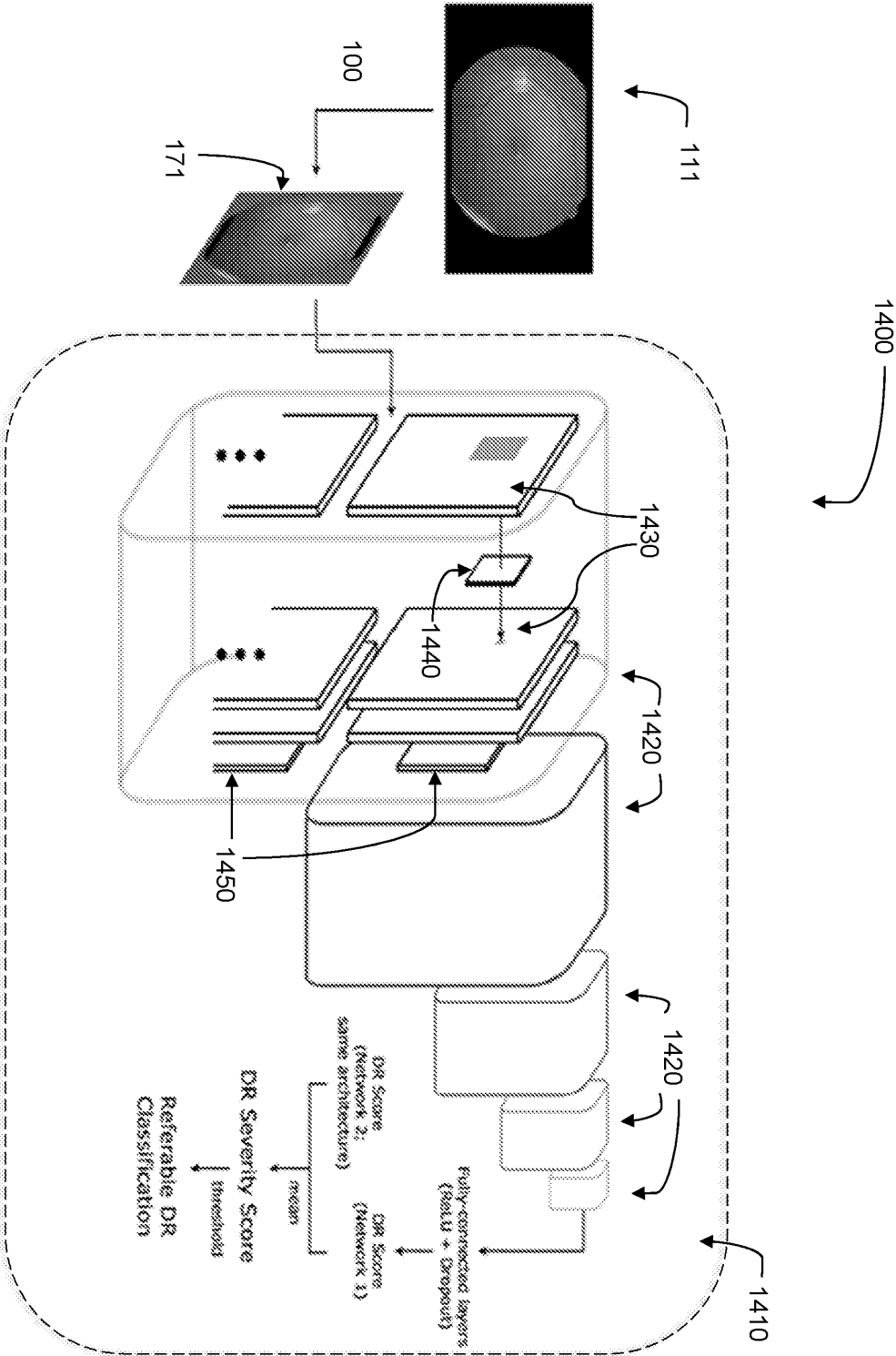


Figure 14

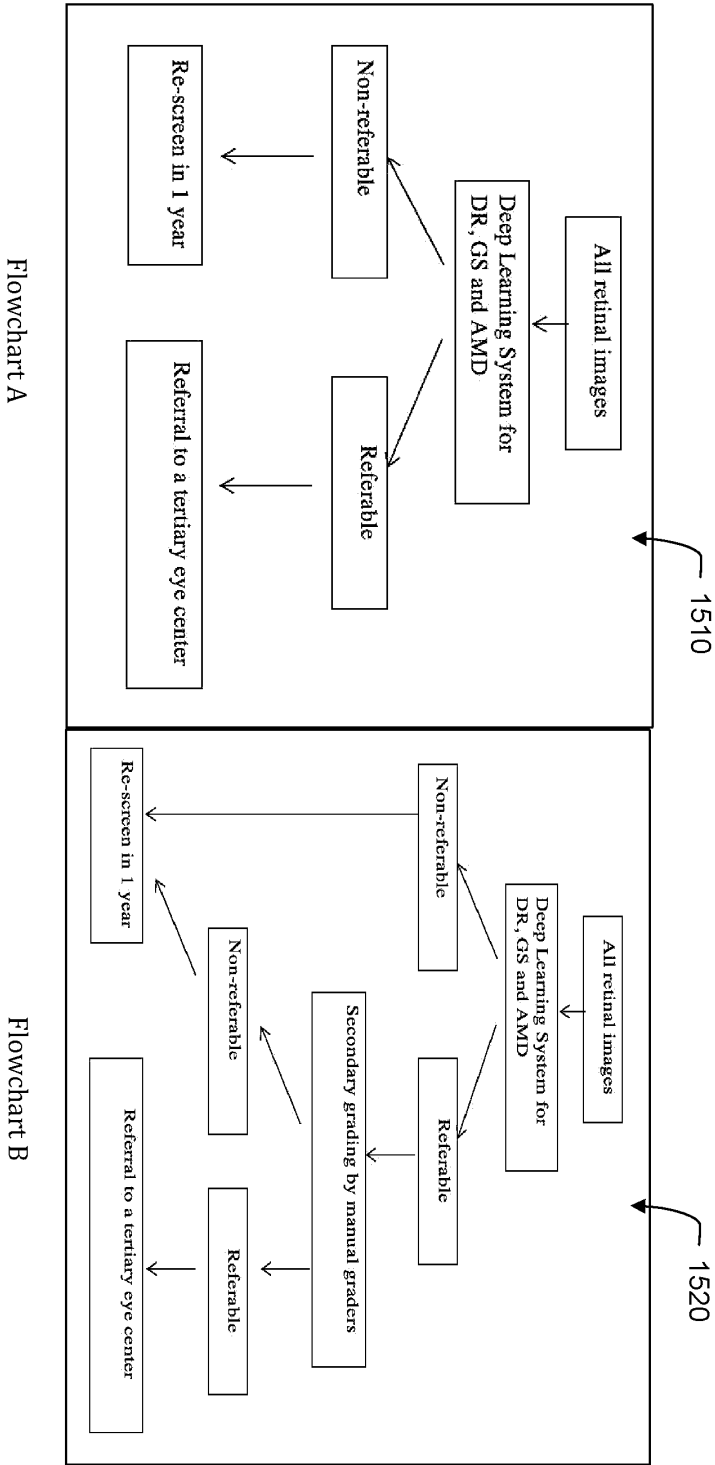


Figure 15

**Table 4: Demographics, diabetes history and systemic risk factors of patients attending national Singapore Diabetic Retinopathy Screening Program in 2014 and 2015 (SIDRP 2014-15, primary validation dataset).**

| <b>Patients' demographics and vascular risk factors</b> | <b>Mean (SD) /<br/>number (%)</b> | <b>% of patient<br/>with data</b> |
|---------------------------------------------------------|-----------------------------------|-----------------------------------|
| Total number of retinal images                          | 71,896                            |                                   |
| Total number of eyes                                    | 35,948                            |                                   |
| Total number of patients                                | 14,880                            |                                   |
| Age (years)                                             | 60.16 (12.19)                     | 56.2                              |
| Gender, Male                                            | 1151 (54.58)                      | 56.2                              |
| Ethnicity                                               |                                   |                                   |
| i. Chinese                                              | 1473 (69.84)                      | 56.2                              |
| ii. Indian                                              | 241 (11.43)                       |                                   |
| iii. Malay                                              | 302 (14.32)                       |                                   |
| iv. Others                                              | 93 (4.41)                         |                                   |
| Systemic risk factors                                   |                                   |                                   |
| 1. BMI (kg/m <sup>2</sup> )                             | 27.22 (4.99)                      | 29.7                              |
| 2. Diabetes duration (years)                            | 3.71 (5.51)                       | 14.7                              |
| 3. Systolic blood pressure (mmHg)                       | 132.05 (17.57)                    | 43.6                              |
| 4. Diastolic blood pressure (mmHg)                      | 72.77 (10.78)                     | 43.6                              |
| 5. HbA1c (%)                                            | 7.54 (1.88)                       | 52.1                              |
| 6. Total cholesterol (mmol/L)                           | 4.65 (1.07)                       | 15.8                              |
| 7. HDL cholesterol (mmol/L)                             | 1.31 (0.35)                       | 15.8                              |
| 8. LDL cholesterol (mmol/L)                             | 2.61 (0.87)                       | 15.4                              |
| 9. Triglycerides (mmol/L)                               | 1.68 (2.22)                       | 15.8                              |
| 10. Creatinine (µmol/L)                                 | 75.34 (36.3)                      | 13.2                              |

**Figure 16**

**Table 5: Primary validation dataset showing the area under curve (AUC), sensitivity, specificity, relative true positive fraction (TPF) and relative false positive fraction (FPF) of deep learning system (DLS) versus trained professional graders in patients with diabetes attended national Singapore diabetic retinopathy screening program (SiDRP 2014-15), with reference to retinal specialist's grading.**

|                         | Diagnostic Performance for Referable DR and VTDR |                                      |                                      |                                        |                                        |
|-------------------------|--------------------------------------------------|--------------------------------------|--------------------------------------|----------------------------------------|----------------------------------------|
|                         | Area under curve<br>(95% CI) <sup>+</sup>        | Sensitivity<br>(95% CI) <sup>^</sup> | Specificity<br>(95% CI) <sup>^</sup> | TPF difference***<br>(95% CI), p-value | FPF difference***<br>(95% CI), p-value |
| <b>1. Referable DR*</b> |                                                  |                                      |                                      |                                        |                                        |
| i. DLS                  | 0.936 (0.925, 0.943)                             | 90.49 (87.25, 92.97)                 | 91.56 (91.01, 92.23)                 | -0.70 (-4.24, 2.85), 0.68              | 7.44 (7.15, 7.72), <0.0001             |
| ii. Graders             | 0.992 (0.987, 0.996)                             | 91.18 (87.97, 93.60)                 | 99.34 (99.24, 99.42)                 | [Reference]                            | [Reference]                            |
| <b>2. VTDR**</b>        |                                                  |                                      |                                      |                                        |                                        |
| i. DLS                  | 0.958 (0.956, 0.961)                             | 100.0 (94.13, 100.00) <sup>#</sup>   | 91.07 (90.70, 91.42)                 | 11.48 (1.84, 21.11), <0.01             | 8.57 (8.27, 8.86), <0.0001             |
| ii. Graders             | 0.997 (0.996, 0.998)                             | 88.52 (75.30, 95.13)                 | 99.63 (99.55, 99.70)                 | [Reference]                            | [Reference]                            |

Eyes were the units of analysis (n=35,948)

DR: Diabetic retinopathy; VTDR: Vision-threatening diabetic retinopathy

^ Asymptotic 95% confidence interval was computed for the logit of each proportion and using the cluster sandwich estimator of standard error to account for possible dependency of eyes within each individual

# Exact Clopper-Pearson left-sided 97.5% confidence interval was calculated due to estimate being at the boundary

+ Cluster-bootstrap biased-corrected 95% confidence interval was computed for each AUC, where individual patients were the bootstrap sampling clusters

\* Referable DR was defined as moderate non-proliferative DR (NPDR), severe (NPDR), proliferative DR (PDR) and 'un-gradable' eye

\*\* VTDR was defined as severe NPDR and PDR

\*\*\* Absolute difference in true positive fraction (TPF) and false positive fraction (FPF) between automated DR screening program and trained graders; TPF is equivalent to sensitivity; FPF is equivalent to 1-specificity

The DME detection rate amongst referable DR eyes = 98.17 (94.37 to 99.42)

**Figure 17**

**Table 6: The area under curve (AUC), sensitivity, specificity, area under curve, relative true positive fraction (TPF) and relative false positive fraction (FPF) of deep learning system (DLS) versus trained professional graders, with reference to retinal specialist's grading in unique SiDRP 14-15 patients.**

|                  | Area under curve<br>(95% CI) <sup>+</sup> | Sensitivity<br>(95% CI) <sup>^</sup> | Specificity<br>(95% CI) <sup>^</sup> | TPF difference***<br>(95% CI), p-value | FPF difference***<br>(95% CI), p-value |
|------------------|-------------------------------------------|--------------------------------------|--------------------------------------|----------------------------------------|----------------------------------------|
| I. Referable DR* |                                           |                                      |                                      |                                        |                                        |
| i. DLS           | 0.879 (0.864, 0.893)                      | 89.56 (85.51, 92.58)                 | 83.49 (82.68, 84.27)                 | 4.71 (-0.16, 9.59), 0.04               | 15.06 (14.25, 15.87), <0.0001          |
| ii. Graders      | 0.917 (0.897-0.937)                       | 84.84 (81.28, 88.51)                 | 98.55 (98.27, 98.79)                 | [Reference]                            | [Reference]                            |
| 2. VTDR**        |                                           |                                      |                                      |                                        |                                        |
| i. DLS           | 0.908 (0.900, 0.915)                      | 100.0 (90.97 to 100.0) <sup>#</sup>  | 81.40 (80.57, 82.22)                 | 10.26 (-0.18, 22.34), 0.04             | 17.68 (16.85, 18.50), <0.0001          |
| ii. Graders      | 0.944 (0.896-0.992)                       | 89.74 (74.77, 96.27)                 | 99.09 (98.86, 99.27)                 | [Reference]                            | [Reference]                            |

16 / 22

**Figure 18**

Patients were the units of analysis (n=8,589)  
DR: Diabetic retinopathy; VTDR: Vision-threatening diabetic retinopathy  
<sup>^</sup> Asymptotic 95% confidence interval was computed for the logit of each proportion  
<sup>#</sup> Exact Clopper-Pearson left-sided 97.5% confidence interval was calculated due to estimate being at the boundary  
<sup>+</sup> Asymptotic 95% confidence interval was computed for each AUC  
<sup>\*</sup> Referable DR was defined as one of the eyes with at least moderate non-proliferative DR (NPDR), severe (NPDR), proliferative DR (PDR) or 'un-gradable'  
<sup>\*\*</sup> VTDR was defined as severe NPDR and PDR  
<sup>\*\*\*</sup> Absolute difference in true positive fraction (TPF) and false positive fraction (FPF) between automated DR screening program and trained graders; TPF is equivalent to sensitivity; FPF is equivalent to 1-specificity  
Eyes rated 'un-gradable' are treated as referable status  
In patients with information in one eye missing, the other eye is used solely to determine referable status  
The DME detection rate amongst referable DR eyes = 97.71 (93.04 to 99.27)

**Table 7: External validation datasets showing the area under curve (AUC), sensitivity, specificity, positive predictive value and negative predictive value of deep learning system (DLS) in detecting referable diabetic retinopathy (DR) and detection rate of vision-threatening DR (VTDR) among populations with diabetes, with reference to trained professional graders**

| Dataset/Ethnicity           |                  | Diagnostic Performance for Referable DR   |                                        |                                      |                                                       |                                                       | VTDR** |
|-----------------------------|------------------|-------------------------------------------|----------------------------------------|--------------------------------------|-------------------------------------------------------|-------------------------------------------------------|--------|
|                             |                  | Area under curve<br>(95% CI) <sup>+</sup> | Sensitivity<br>(95% CI) <sup>^</sup>   | Specificity<br>(95% CI) <sup>^</sup> | Positive<br>predictive<br>value (95% CI) <sup>^</sup> | Negative<br>predictive<br>value (95% CI) <sup>^</sup> |        |
| <i>i) Community-based</i>   |                  |                                           |                                        |                                      |                                                       |                                                       |        |
| 1. Chinese                  | Guangdong, China | 0.949<br>(0.943, 0.955)                   | 98.65<br>(97.66, 99.3)                 | 81.62<br>(80.69, 82.53)              | 40.94<br>(38.85, 43.06)                               | 99.79<br>(99.63, 99.89)                               | 100.0  |
| <i>ii) Population-based</i> |                  |                                           |                                        |                                      |                                                       |                                                       |        |
| 2. Chinese                  | Singapore        | 0.919<br>(0.900, 0.942)                   | 100.00<br>(92.52, 100.00) <sup>#</sup> | 76.29<br>(72.69, 79.55)              | 24.37<br>(18.61, 31.23)                               | 100.00<br>(99.46, 100.00) <sup>#</sup>                | 100.0  |
| 3. Malay                    | Singapore        | 0.889<br>(0.863, 0.908)                   | 97.14<br>(92.52, 98.94)                | 82.03<br>(79.37, 84.42)              | 35.79<br>(29.91, 42.13)                               | 99.64<br>(99.05, 99.87)                               | 92.6   |
| 4. Indian                   | Singapore        | 0.917<br>(0.899, 0.933)                   | 99.32<br>(95.14, 99.91)                | 73.27<br>(70.86, 75.54)              | 20.95<br>(17.32, 25.11)                               | 99.93<br>(99.53, 99.99)                               | 100.0  |
| 5. Chinese                  | Beijing, China   | 0.929<br>(0.903, 0.955)                   | 94.44<br>(72.71, 99.86)                | 88.51<br>(85.41, 91.16)              | 22.67<br>(13.79, 33.79)                               | 99.78<br>(98.76, 99.99)                               | 100.0  |
| 6. African Americans        | Los Angeles, USA | 0.980<br>(0.971, 0.989)                   | 98.8<br>(93.47, 99.97)                 | 86.51<br>(84.05, 88.72)              | 41.41<br>(34.48, 48.61)                               | 99.87<br>(99.25, 100)                                 | 97.0   |
| <i>iii) Clinic-based</i>    |                  |                                           |                                        |                                      |                                                       |                                                       |        |

**Figure 19A**

|              |                      |                         |                         |                         |                         |                         |       |
|--------------|----------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------|
| 7. Caucasian | Melbourne, Australia | 0.983<br>(0.972, 0.991) | 98.94<br>(97.47, 99.56) | 92.22<br>(89.52, 94.27) | 91.59<br>(88.68, 93.85) | 99.03<br>(97.68, 99.60) | 100.0 |
| 8. Hispanic  | Mexico               | 0.950<br>(0.934, 0.966) | 91.76<br>(88.44, 94.37) | 84.78<br>(80.39, 88.53) | 87.21<br>(83.44, 90.38) | 90.1<br>(86.17, 93.22)  | 98.8  |
| 9. Chinese   | Hong Kong            | 0.948<br>(0.921, 0.972) | 99.31<br>(97.25, 99.83) | 83.14<br>(77.93, 87.31) | 83.43<br>(78.24, 87.58) | 99.29<br>(97.17, 99.83) | 98.2  |
| 10. Chinese  | Hong Kong            | 0.964<br>(0.958, 0.97)  | 100<br>(98.95, 100) #   | 81.32<br>(79.99, 82.6)  | 34.93<br>(31.98, 37.96) | 100<br>(99.87, 100) #   | 100   |

^ Asymptotic 95% confidence interval was computed for the logit of each proportion and using the cluster sandwich estimator of standard error to account for possible dependency of eyes within each individual

# Exact Clopper-Pearson left-sided 97.5% confidence interval was calculated due to estimate being at the boundary

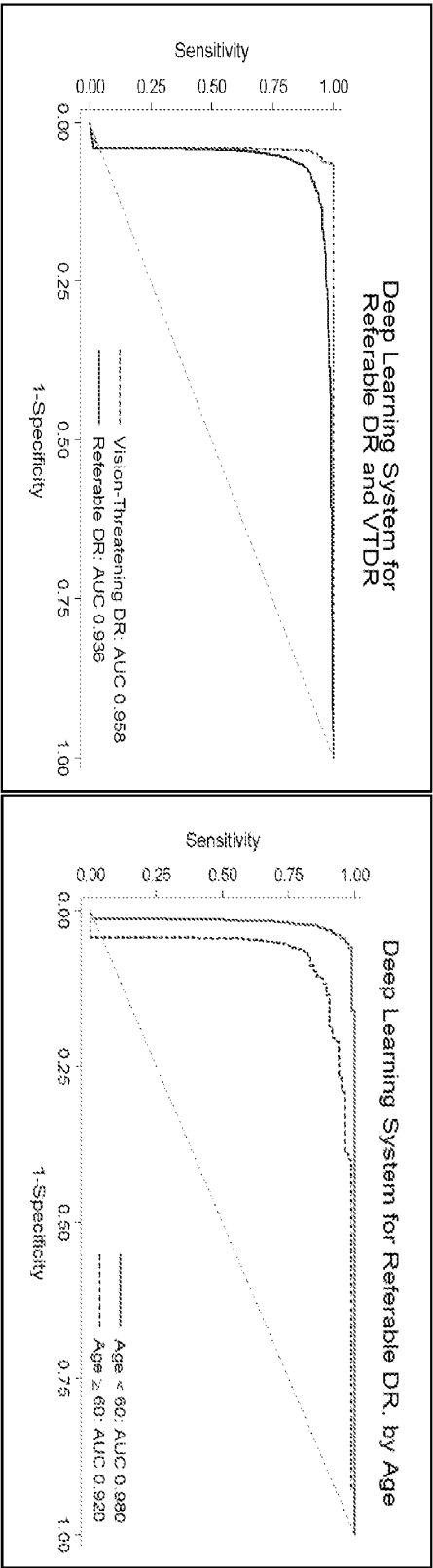
+ Cluster-bootstrap biased-corrected 95% confidence interval was computed for each AUC, where individual patients were the bootstrap sampling clusters

\* Referable DR was defined as moderate non-proliferative DR (NPDR), severe (NPDR), proliferative DR (PDR) and un-gradable images

\*\* VTDR was defined as severe NPDR and PDR. The detection rate is the proportion of VTDR detected using the same threshold set for referable DR. The AUC for Dataset 1 to 10 are 0.926 (dataset 1), 0.959 (dataset 2), 0.875 (dataset 3), 0.912 (dataset 4), 0.914 (dataset 5), 0.952 (dataset 6), 0.936 (dataset 7), 0.822 (dataset 8), 0.906 (dataset 9) and 0.958 (dataset 10), respectively.

Figure 19B

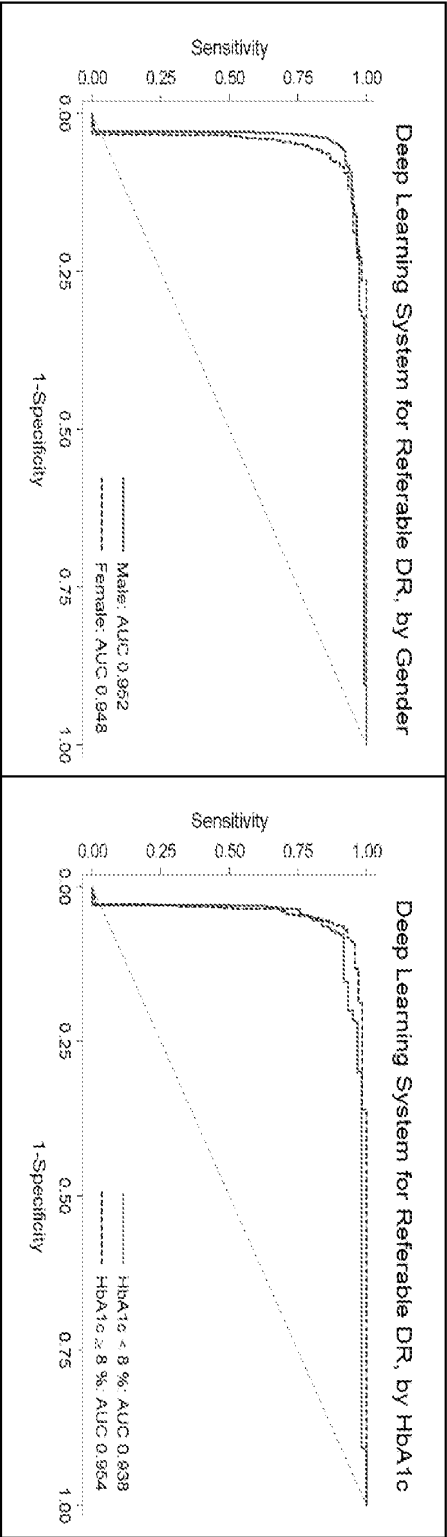
- The receivers' operating curve (ROC) and area under curve (AUC) of deep learning system (DLS) for detection of:
- (A) Referable diabetic retinopathy (DR) and vision-threatening diabetic retinopathy (VTDR) in the national Singapore diabetic retinopathy screening program (SIDRP 2014-15);
  - (B) Referable DR stratified by age, gender and glycemic control in SIDRP 2014-15
  - (C) Referable DR and VTDR in 10 populations of different ethnicities – Chinese (Guangdong, China), Chinese (Singapore), Malay (Singapore), Indian (Singapore), Chinese (Beijing, China), African American (Los Angeles, USA), Caucasian (Melbourne, Australia), Hispanic (Mexico), Chinese (Hong Kong) and Chinese (Hong Kong).



Graph A

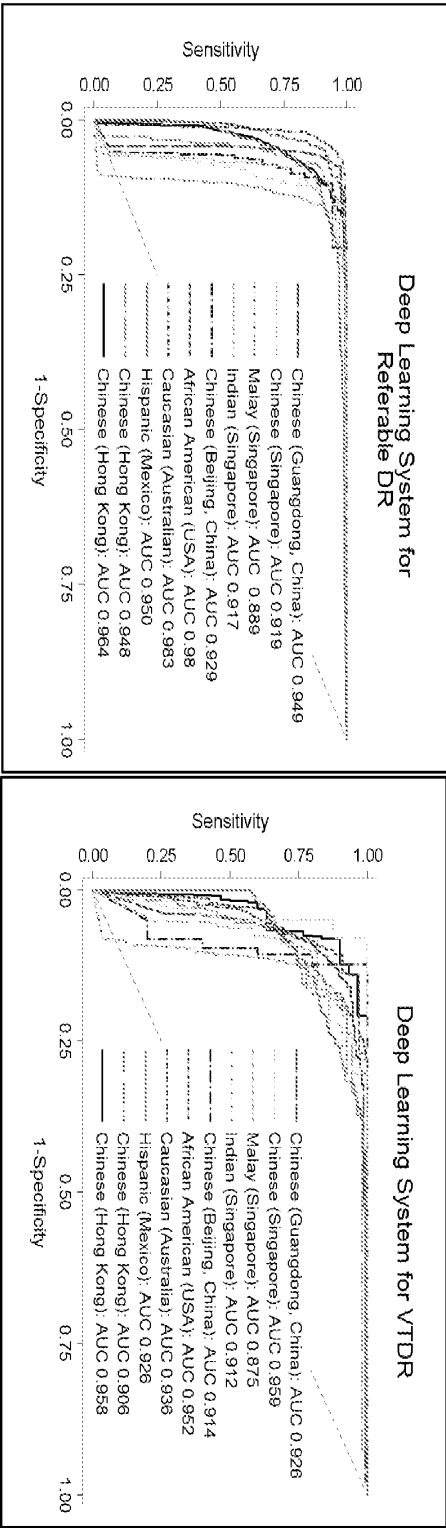
Graph B1

Figure 20A



Graph B2

Graph B3



Graph C1

Graph C2

Figure 20B

Eyes are the units of analysis

Cluster-bootstrap biased-corrected 95% confidence interval (CI) was computed for each AUC, where individual patients were the bootstrap sampling clusters

Referable diabetic retinopathy (DR): Defined as moderate non-proliferative DR (NPDR), severe NPDR, proliferative DR and diabetic macular oedema

Vision-threatening diabetic retinopathy: Defined as severe NPDR and PDR

A: SIDRP 2014-15 cohort (with reference to a retinal specialist)

i. Referable DR: 0.936 (0.925, 0.943); VTDR: 0.958 (0.956, 0.961)

B: SIDRP 2014-15 cohort stratified by age, gender and HbA1c (with reference to a retinal specialist)

i. Age <60 years: 0.980 (0.975, 0.984); Age ≥ 60: 0.920 (0.889, 0.940)

ii. Male: 0.952 (0.925, 0.963); Female: 0.948 (0.933, 0.956); 3)

iii. HbA1c <8%: 0.938 (0.892, 0.958); HbA1c ≥ 8%: 0.954 (0.942, 0.964)

C: Referable DR and VTDR in 10 populations with different ethnicities (with reference to professional graders)

i. Referable DR – Chinese (Guangdong, China): 0.949 (0.943, 0.955); Chinese (Singapore): 0.919 (0.900, 0.942); Malay (Singapore): 0.889 (0.863, 0.908); Indian (Singapore): 0.917 (0.899, 0.933); Chinese (Beijing, China): 0.929 (0.903, 0.955); African American (USA): 0.980 (0.971, 0.989); Caucasian (Melbourne, Australia): 0.983 (0.972, 0.991); Hispanic (Mexico): 0.950 (0.934, 0.956); Chinese (Hong Kong): 0.948 (0.921, 0.972); Chinese (Hong Kong): 0.964 (0.958, 0.970)

ii. VTDR – Chinese (Guangdong, China): 0.926 (0.910, 0.943); Chinese (Singapore): 0.959 (0.943, 0.974); Malay (Singapore): 0.875 (0.845, 0.894); Indian (Singapore): 0.912 (0.855, 0.936); Chinese (Beijing, China): 0.914 (0.874, 0.953); African American (USA): 0.952 (0.931, 0.974); Caucasian (Melbourne, Australia): 0.936 (0.915, 0.950); Hispanic (Mexico): 0.822 (0.766, 0.864); Chinese (Hong Kong): 0.906 (0.858, 0.938) and; Chinese (Hong Kong): 0.958 (0.940, 0.976).

Figure 20C

**Table 8: Primary validation dataset showing the area under curve (AUC), sensitivity, specificity, positive predictive value and negative predictive value of deep learning system (DLS) in detecting referable glaucoma suspect (GS) and referable AMD (AMD) among patients with diabetes attending national Singapore diabetic retinopathy screening program (SIDRP 2014-15), with reference to a retinal and glaucoma specialists.**

|                      | Diagnostic Performance for Referable GS and Referable AMD |                                      |                                      |                                                    |                                                    |
|----------------------|-----------------------------------------------------------|--------------------------------------|--------------------------------------|----------------------------------------------------|----------------------------------------------------|
|                      | Area under curve<br>(95% CI) <sup>+</sup>                 | Sensitivity<br>(95% CI) <sup>^</sup> | Specificity<br>(95% CI) <sup>^</sup> | Positive predictive<br>value (95% CI) <sup>^</sup> | Negative predictive<br>value (95% CI) <sup>^</sup> |
| <i>Referable GS</i>  | 0.942<br>(0.929,0.954)                                    | 96.40<br>(81.65,99.91)               | 87.17<br>(86.81,87.52)               | 0.61<br>(0.4,0.88)                                 | 100<br>(99.98,100) <sup>#</sup>                    |
| <i>Referable AMD</i> | 0.931<br>(0.928,0.935)                                    | 93.17<br>(91.14,99.75)               | 88.65<br>(88.29,89.01)               | 17.46<br>(16.3,18.7)                               | 99.8<br>(99.74,100) <sup>#</sup>                   |

22 / 22

**Figure 21**

GS: glaucoma suspect; AMD: age-related macular degeneration  
Referable GS: defined as vertical cup/disc diameter ratio of  $\geq 0.8$ , focal thinning or notching of the neuroretinal rim, optic disc haemorrhages, or localized retinal nerve fibre layer defects.  
Referable AMD: defined as the presence of intermediate AMD (numerous intermediate drusens, 1 large drusen  $>125\mu\text{m}$ ) and/or advanced AMD - geographic atrophy (GA) or neovascular AMD, using the age-related eye disease study (AREDS) grading system.<sup>52</sup>  
Repeats from SIDRP 14-15 were excluded from the analysis  
<sup>^</sup> Asymptotic 95% confidence interval was computed for the logit of each proportion and using the cluster sandwich estimator of standard error to account for possible dependency of eyes within each individual  
<sup>#</sup> Exact Clopper-Pearson left-sided 97.5% confidence interval was calculated due to estimate being at the boundary  
<sup>+</sup> Cluster-bootstrap biased-corrected 95% confidence interval was computed for each AUC, where individual patients were the bootstrap sampling clusters