

Multimedia Appendix 7. Full table of sensitivity, specificity, accuracy, F1 score and MCC for all symptom checkers (COVID-19 positive defined by “medium risk or “high risk” for non binary symptom checkers)

| Symptom checker | sensitivity | specificity | accuracy | F1 score | MCC |
|-------------------------|-------------|-------------|----------|----------|------|
| Ada | 0.36 | 0.67 | 0.51 | 0.42 | 0.03 |
| Apple | 0.98 | 0.17 | 0.57 | 0.70 | 0.25 |
| Babylon | 0.90 | 0.33 | 0.62 | 0.70 | 0.29 |
| CDC | 0.94 | 0.29 | 0.61 | 0.71 | 0.30 |
| Cleveland Clinic | 0.98 | 0.41 | 0.69 | 0.76 | 0.47 |
| Docyet | 0.16 | 1.00 | 0.58 | 0.27 | 0.29 |
| Infermedica | 0.78 | 0.83 | 0.80 | 0.80 | 0.61 |
| Providence | 0.98 | 0.38 | 0.68 | 0.75 | 0.45 |
| SF-COS | 0.90 | 0.66 | 0.78 | 0.80 | 0.58 |
| SF-DIST | 0.52 | 0.70 | 0.61 | 0.57 | 0.22 |
| Symptoma | 0.94 | 0.88 | 0.91 | 0.91 | 0.83 |
| Your.MD | 0.94 | 0.32 | 0.63 | 0.72 | 0.33 |