

Predicting Which Patients Are at Risk for Clinical Deterioration in COVID-19: A Review of the Current Models in Use

Michael Shapiro MD^{1,3}, Yarden Yavne MD^{1,3}, and Daniel Shepshelovich MD^{2,3}

Departments of ¹Internal Medicine T and ²Internal Medicine D, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel
³Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

ABSTRACT The ongoing coronavirus disease 2019 (COVID-19) pandemic has led to more than 200 million infected cases and 4.6 million deaths worldwide, and the numbers continue to grow. The disease presentation varies, and while most patients will present with a mild disease course, 5% will eventually develop significant respiratory failure, some despite initially presenting with mild symptoms. Early detection of patients at risk for deterioration is crucial for decisions regarding hospitalization, monitoring, timing, and extent of treatment.

IMAJ 2022; 24: 699–704

KEY WORDS: coronavirus disease 2019 (COVID-19), machine learning, prediction models, risk factors, severe disease

(see editorial page 705)

Numerous studies examining risk factors have found that age above 65 years, obesity, history of smoking, and other co-morbidities as well as various laboratory markers such as lymphopenia, elevated D-dimer and troponin are associated with a more severe disease course among COVID-19 patients. Models comprising various combinations of these risk factors, together with vital signs, laboratory tests, and radiographic features have been developed to construct a reliable prediction model for patient deterioration. Several studies have attempted to develop novel algorithms specifically targeting coronavirus disease 2019 (COVID-19) patients, some of them with the help of machine learning, while others examined the applicability of re-purposing already available and validated prediction models from other diseases for COVID-19.

Despite the extensive research, currently there is no formal recommendation for the use of a single model and physicians must evaluate the suitability of the available options for each

case separately. We outlined the leading models in this field and discussed the tools that are available for critical assessment of current and future models.

The onset of the COVID-19 pandemic first began in November 2019 in Wuhan, China. By September 2021, more than 226 million cases were detected worldwide with more than 4.6 million deaths [1]. The disease course varies among patients and while the majority (80%) present with mild symptoms such as cough and fever, approximately 14% develop severe symptoms such as shortness of breath, hypoxia, and diffuse lung involvement. Another 5% ultimately progress to respiratory, hemodynamic, and multi-organ failure [2]. Some of the patients presenting initially with mild symptoms later deteriorate approximately a week after symptom onset [3]. Early detection of these patients can assist physicians to provide individualized decisions regarding the need for hospitalization, monitoring, and treatment suitable for each patient.

Since the emergence of the COVID-19 pandemic, multiple studies have detected the relevant risk factors for prediction of a severe disease course. Several prediction tools have been designed to assist physicians. The physician has the responsibility to critically examine the

available methods and subsequently decide whether the research and methodology behind these methods justifies their use. In this review,

we evaluated and discussed various existing models for prognosis prediction of COVID-19 patients. For the sake of brevity, we emphasized models that were published in major journals, most of which were externally validated.

To assess the available models, it was important to differentiate between two major types of studies. The first type includes studies that evaluated risk factors determining whether the patient's characteristics (such as age, symptoms, co-morbidities,

EARLY DETECTION OF COVID-19 PATIENTS WHO ARE AT RISK FOR DETERIORATION CAN ASSIST PHYSICIANS IN MAKING DECISIONS REGARDING THE NEED FOR HOSPITALIZATION, MONITORING, AND TREATMENT

or biomarkers) are independently associated with clinical outcomes (such as hospitalization, disease severity, and mortality) [4]. The second type refers to studies that suggested a model for the prediction of a clinical outcome for each patient by combining the previously determined risk factors [5].

STUDIES EVALUATING RISK FACTORS

Several studies that examined the association between various demographic factors and mortality and/or disease severity among COVID-19 patients found that age > 65 years, male sex, obesity (body mass index > 30 kg/m²), and smoking history were significant risk factors [6,7]. Moreover, patients with co-morbidities such as chronic obstructive pulmonary disease, chronic heart failure, ischemic heart disease, cancer, diabetes, chronic kidney disease, and hypertension were at a higher risk of developing severe disease [6,7]. While many of the initial large studies regarding risk factors were performed on Chinese patients, subsequent studies have shown similar results in Europe, the United States, and other countries [8].

As the pandemic progressed, evidence surfaced regarding the correlation between laboratory results and a more severe disease. Neutropenia; lymphopenia [9]; increased neutrophil-to-lymphocyte ratio [10]; and elevated lactate dehydrogenase (LDH), C-reactive protein (CRP), D-dimer, and troponin levels [11] were found to be associated with severe disease. It is important to note that these studies were retrospective in nature. Many collected data regarding the laboratory results and the patient status concurrently. In addition, many of these studies did not differentiate between patients who were already severely ill at the beginning of the follow-up period from those who presented with mild symptoms.

It is important to note that the immunization status of a patient, including the time from the last immunization, has been shown in multiple large cohort and randomized control trial to be a significant independent protective factor for both severe disease and mortality [12,13].

STUDIES DEVELOPING A PREDICTION MODEL

To evaluate the efficiency and applicability of a prediction model, it is important to employ statistical measures such as the positive predictive value (PPV), negative predictive value (NPV), accuracy, sensitivity, and specificity. Due to the trade-off between these measures, a different cut-off can improve one value at the expense of another.

THE RECEIVER OPERATING CHARACTERISTIC-AREA UNDER THE CURVE MEASURE

The receiver operating characteristic-area under the curve measure (ROC-AUC) measurement is a metric developed to eval-

uate the difference between sensitivity and specificity and is equal to the AUC of the sensitivity of the model as a function of the false positive rate (equals 1 minus specificity). The value provides a good measure of the discriminatory capacity of classification models, ranging from 0.5 (equal to random) to 1 (a perfect predictive model), with 0.8 generally regarded as a good value [14].

The prediction models presented here are divided into two categories: established models that were validated to predict the risk for morbidity and mortality in a similar disease (e.g., pneumonia) and were repurposed for COVID-19 patients, and new models that were designed a priori for

the evaluation of COVID-19 patients. The major asset of using established models is their extensive use in clinical practice, whereas newer models have the advantage of being constructed specifically for COVID-19, a novel disease with a unique disease course amongst respiratory infections.

It is important to note that current models were built and tested among unvaccinated patients and thus might need to be adjusted for use among vaccinated patients.

ESTABLISHED MODELS

- **CURB-65** is a common score for evaluating the mortality risk of patients presenting with community acquired pneumonia in the emergency department (ED). It has been validated extensively, recommending hospitalization for patients with a score of two or more [15]. Several studies demonstrated an AUC of about 0.8 for COVID-19 mortality and severe disease but a poor sensitivity of 65% using the original cutoff [16,17] (compared to 91% sensitivity for severe pneumonia), while another study showed a much lower AUC [18].
- **A-DROP** is a variation of CURB-65 score which replaces the parameter of respiratory rate with O₂ saturation < 90% [19]. The score showed promising results for predicting 30-day mortality among hospitalized patients with a sensitivity of 80% and specificity of 86% in one study [20] but was not assessed further.
- **Pneumonia severity index (PSI)** is another common score for the evaluation of community acquired pneumonia patients in the ED, which is based on co-morbidities, laboratory tests, and vital signs. Patients in category 3 or higher require hospitalization [21]. The score outperformed CURB-65 for predicting mortality in two studies, both in terms of AUC and sensitivity [16,20].
- **National Early Warning Score 2 (NEWS2)** is a popular model for severity assessment of hospitalized patients in the United Kingdom. The score was recommended by the Royal College of Physicians for

use among COVID-19 patients at the beginning of the pandemic. Several studies showed that the model tends to underestimate the risk for severe disease [22] and mortality [23] among COVID-19 patients.

- **Q-SOFA** is a popular screening score for sepsis [24] due to its prognosis prediction in smaller studies, but it showed consistently poorer results for mortality prediction with AUC of 0.6–0.7 and low sensitivity [23,25].

MODELS DESIGNED SPECIFICALLY FOR COVID-19 PATIENTS

- **CALL score** is a small but highly cited study, published during the first months of the pandemic. It was developed as a model among 208 patients and included lymphocyte number, age, co-morbidities, and LDH levels as parameters. The model showed good results on internal validation, with an AUC of 0.91 [26]. However, when examined independently by an Italian group, the performance worsened with an AUC of 0.6. Therefore, we do not recommend the use of this score [27]
- **COVID-GRAM** is a model developed to predict severe disease or death among a group of 1600 Chinese hospitalized patients and validated in an additional study of 700 patients. The model included 19 parameters including X-ray results, vital signs, rash, hemoptysis, loss of consciousness, LDH, and bilirubin and creatinine levels. The model showed an impressive result with an AUC 0.88. It is important to note that in contrast to many studies that included a population with high rate of severely ill patients (20%), this study population closely resembled the general population, with mortality rate of 1% [28]. A validation of this score among 214 Spanish intensive care patients showed an AUC of 0.72. While the authors of the original article did not examine a cutoff point for the score, the validation article showed that using a cutoff score of 89 provided adequate sensitivity but lacked specificity (32%), limiting the score usefulness [29].
- **QCOVID** tool is the result of a study performed by the University of Oxford with the goal of creating a model capable of evaluating the risk for hospital admission and mortality among non-hospitalized COVID-19 patients [30]. It includes age, socioeconomic parameters, smoking status, and body mass index, as well as co-morbidities. The tool was developed and internally and externally validated on millions of British patients [31], reporting a sensitivity

of 76% in the original study and up to 70% in the validation cohort while keeping a reasonable false positive rate. The use of the Townsend Deprivation Index, a socioeconomic score specific for British citizens as an important parameter, limits the use of the tool outside the United Kingdom.

- **Quick COVID-19 Severity Index** is a unique model that predicts the risk for respiratory failure within 24 hours of hospital admission. The quick version, including respiratory rate, pulse oximetry, and the required oxygen supplementation flow, had an AUC score of 0.81. The score of 3 had a sensitivity and specificity of 79%; however, the authors of the original article concluded that there were few patients in the validation cohort above the cutoff, thus these metrics might not be accurate [18]. When performance was evaluated in a validation study examining mortality prediction, the model performed worse, achieving an AUC score of 0.71, and was inferior to the CURB-65 model (AUC 0.78) [32].
- **VACO Index** is a model developed to predict the 30-day all-cause mortality among inpatients and outpatients, developed and tested among 13,000 U.S. veterans [33]. A validation study demonstrated the model's discriminatory power (AUC 0.82) for the entire population similar to the results of the original study, but among patients older than 65 years of age, which are the patients most at risk in COVID-19, performance dropped (AUC 0.69) [34].

**A CLINICALLY USEFUL MODEL SHOULD BE DESIGNED
AND EXTERNALLY VALIDATED FOR THE RELEVANT
POPULATION, POSSESS AN ACCEPTABLE DISCRIMINATORY
POWER, AND SUGGEST A CUTOFF FOR WHICH SENSITIVITY
AND SPECIFICITY ARE AVAILABLE**

MACHINE LEARNING MODELS DESIGNED FOR COVID-19 PATIENTS

- **Yan L. Biomarkers based model** is a promising model published in August 2020, which identified LDH, CRP, and lymphocytes as key features and created a machine-learning model (XGBoost based) for mortality prediction on hospital admission, achieving an AUC of 0.95 [35]. However, an external validation study demonstrated that the model performs poorly with specificity of 26% (compared to 96% in the original study), thus it has limited clinical use [36], emphasizing the importance of external validation.
- **Random Forest algorithm (Denmark and UK) and LASSO (Korean) models** are two large studies including thousands of patients that examined machine learning methods for mortality prediction among hospitalized patients and achieved impressive AUC scores of above 0.96 for mortality prediction but performed worse when tested among a different pop-

Table 1. Summary of prediction models

Model	Designated population, cohort-training (instrument)	Original performance (external validation performance)	Comments
Established models			
CURB-65	ED patients	AUC 0.80 Sensitivity 65% Specificity 91%	Large performance variance across studies
A-DROP	Hospitalized patients	AUC 0.87 Sensitivity 80% Specificity 86%	Relatively untested for COVID-19 patients
Pneumonia severity index	ED patients	AUC 0.85 Sensitivity 77% Specificity 81%	Large performance variance across studies
National Early Warning Score 2	Hospitalized patients	AUC 0.70 Sensitivity 78% Specificity 48%	Mainly used in the UK in the early stages of the pandemic but later replaced by other models
Q-SOFA	Hospitalized patients	AUC 0.60–0.7 Sensitivity 42% Specificity 84%	Failed to classify severe COVID-19 patient as having sepsis
Models designed specifically for COVID-19 patients			
CALL score (China)	Hospitalized patients Cohort of 208 patients	AUC 0.91 (0.60) Sensitivity and Specificity not provided	External validation showed poor results
COVID-GRAM (China)	Hospitalized patients Cohort of 1600 (700)	AUC 0.88 (0.72) Sensitivity and Specificity not provided	Cohort had only 1% mortality, resembling the general population
QCOVID (Britain)	General population Cohort of 6.18 million (2.17 million)	AUC 0.88 (0.72) Sensitivity and Specificity not provided	The use of the Townsend Score, specific for Britain, limits international use
Quick COVID-19 Severity Index (USA)	Hospitalized patients Predicting respiratory failure Cohort 932 (240)	AUC 0.81 Sensitivity 79% Specificity 79%	Inferior to CURB-65 for predicting mortality in an external validation study
VACO Index (USA)	Hospitalized and outpatients Cohort 3681 (2151) The text says 13,000	AUC 0.84 Sensitivity and Specificity not provided	Validation study showed poor performance among patients age > 65 years of age
Machine learning models designed for COVID-19 patients			
Yan L. Biomarkers based model (China)	Hospitalized patients 375 (110)	AUC 0.95 Sensitivity 94% (95%) Specificity 96% (26%)	Poor performance on external validation
Random Forest algorithm (Denmark-UK study)	General population Denmark -3944 (UK -1650)	AUC 0.9 (0.74) Sensitivity and Specificity not provided	When tested among UK patients, performed significantly worse
LASSO algorithm (Korean study)	General population 7165 (3071)	AUC 0.96 Sensitivity 91% Specificity 91%	The model was not validated by an additional study nor compared to other models
Image-analysis-based models (CT analysis)	Hospitalized patients 2778 (208)	AUC 0.96 Sensitivity 94% Specificity 91%	The model was not validated by an additional study nor compared to other models

AUC = area under the curve, ED = emergency department, UK = United Kingdom, USA = United States of America

ulation [37,38]. Unfortunately, neither compared their performance to other validated scores, making it impossible to evaluate the added benefit of these methods.

- **Image-analysis-based models (CT analysis)** are two major models that were based on large cohorts of thousands of patients and used machine-learning to extract radiological features from computed tomography scans to evaluate the prognosis of COVID-19 patients. Both studies combined the radiological features with patient vitals and laboratory results to create a prediction model, demonstrating an impressive AUC score of approximately 0.96 [39,40]. However, as with previously mentioned machine learning models, their performance was not compared to other non-image-based scores, hindering the ability to evaluate their added benefit.

CONCLUSIONS

An external validation of a model important for assuring the generalizability of the model. This finding stems from the risk of overfitting, where a model is highly suited for the cohort of patients used to develop it but performs much worse for other populations. In addition to external validation, we recommend considering several important points when assessing any prediction model:

- For the model to be useful in the clinical setting, the study must specify a cut-off score that differentiates high-risk from low-risk patients and calculates the sensitivity and specificity of that cutoff.
- The model should have an acceptable discriminatory power for differentiating between low- and high-risk patients, as represented by a high AUC score.
- The model should be used only for the population among which it was constructed and tested (e.g., general population, patients admitted to the emergency department) until further studies have validated its applicability for other populations.

Correspondence

Dr. M. Shapiro

Dept of Internal Medicine T, Tel Aviv Sourasky Medical Center, Tel Aviv 6423906, Israel

email: mikehpg@gmail.com

References

1. Anon. WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data. (This reference is not complete. Is there a URL? What date did you access the site?)
2. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* 2020; 323 (13): 1239-42.
3. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel

coronavirus in Wuhan, China. *Lancet* 2020; 395 (10223): 497-506.

4. Bouwmeester W, Zuithoff NPA, Mallett S, et al. Reporting and methods in clinical prediction research: a systematic review. *MacLeod MR, ed. PLoS Med* 2012; 9 (5): e1001221.
5. Steyerberg EW, Moons KGM, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research. *PLoS Med* 2013; 10 (2): e1001381.
6. Zheng Z, Peng F, Xu B. Risk factors of critical & mortal COVID-19 cases: a systematic literature review and meta-analysis. *J Infect* 2020; 81: 16-25.
7. Gupta S, Hayek SS, Wang W, et al; STOP-COVID investigators. factors associated with death in critically ill patients with coronavirus disease 2019 in the US. *JAMA Intern Med* 2020; 180 (11): 1436-47.
8. Wingert A, Pillay J, Gates M, et al. Risk factors for severity of COVID-19: a rapid review to inform vaccine prioritisation in Canada. *BMJ Open* 2021; 11 (5): e044684.
9. Guan W, Ni Z, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med* 2020; 382 (18): 1708-20.
10. Lagunas-Rangel FA. Neutrophil-to-lymphocyte ratio and lymphocyte-to-C-reactive protein ratio in patients with severe coronavirus disease 2019 (COVID-19): A meta-analysis. *J Med Virol* 2020; 92 (10): 1733-4.
11. Terpos E, Ntanas-Stathopoulos I, Elalamy I, et al. Hematological findings and complications of COVID-19. *Am J Hematol* 2020; 95 (7): 834-47.
12. Bar-On YM, Goldberg Y, Mandel M, Bodenheimer O, Freedman L, Kalkstein N, Mizrahi B, Alroy-Preis S, Ash N, Milo R, Huppert A, et al. Protection of BNT162b2 vaccine booster against Covid-19 in Israel. *N Engl J Med* 2021; 385 (15): 1393-400.
13. Thomas SJ, Edson D, Moreira J, Kitchin N, et al; C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine through 6 months. *N Engl J Med* 2021; 385 (19): 1761-73.
14. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol* 2010; 5 (9): 1315-6.
15. Lim WS, Van Der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax* 2003; 58 (5): 377-82.
16. Satici C, Demirkol MA, Sargin Altunok E, et al. Performance of pneumonia severity index and CURB-65 in predicting 30-day mortality in patients with COVID-19. *Int J Infect Dis* 2020; 98: 84-9.
17. Luo M, Liu J, Jiang W, Yue S, Liu H, Wei S. IL-6 and CD8+ T cell counts combined are an early predictor of in-hospital mortality of patients with COVID-19. *JCI Insight* 2020; 5 (13): e139024.
18. Haimovich AD, Ravindra NG, Stoytchev S, et al. Development and validation of the quick COVID-19 severity index: a prognostic tool for early clinical decompensation. *Ann Emerg Med* 2020; 76 (4): 442-53.
19. Shindo Y, Sato S, Maruyama E, et al. Comparison of severity scoring systems A-DROP and CURB-65 for community-acquired pneumonia. *Respirology* 2008; 13 (5): 731-5.
20. Fan G, Tu C, Zhou F, et al. Comparison of severity scores for COVID-19 patients with pneumonia: a retrospective study. *Eur Respir J* 2020; 56 (3): 2002113.
21. Arnold FW, Ramirez JA, McDonald LC, Xia EL. Hospitalization for community-acquired pneumonia: The pneumonia severity index vs clinical judgment. *Chest* 2003; 124 (1): 121-4.
22. Carr E, Bendayan R, Bean D, et al. Evaluation and improvement of the National Early Warning Score (NEWS2) for COVID-19: a multi-hospital study. *BMC Med* 2021 191 2021; 19 (1): 1-16.
23. Bradley P, Frost F, Tharmaratnam K, Wootton DG. Utility of established prognostic scores in COVID-19 hospital admissions: multicentre prospective evaluation of CURB-65, NEWS2 and qSOFA. *BMJ Open Respir Res* 2020; 7 (1): e000729.
24. Raith EP, Udy AA, Bailey M, et al. Prognostic accuracy of the SOFA score, SIRS criteria, and qSOFA score for in-hospital mortality among adults with suspected infection admitted to the intensive care unit. *JAMA* 2017; 317 (3): 290-300.
25. Ferreira M, Blin T, Collierand N, et al. Critically ill SARS-CoV-2-infected patients are not stratified as sepsis by the qSOFA. *Ann Intensive Care* 2020 101 2020; 10 (1): 1-3.
26. Ji D, Zhang D, Xu J, et al. Prediction for progression risk in patients with COVID-19 pneumonia: the CALL score. *Clin Infect Dis* 2020; 71 (6): 1393-9.
27. Grifoni E, Valoriani A, Cei F, et al. The CALL score for predicting outcomes in patients with COVID-19. *Clin Infect Dis* 2021; 72 (1): 182-3.
28. Liang W, Liang H, Ou L, et al. Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients with COVID-19. *JAMA Intern Med* 2020; 180 (8): 1081-9.

29. Moreno-Pérez Ó, Andrés M, León-Ramírez JM, et al. The COVID-GRAM Tool for patients hospitalized With COVID-19 in Europe. *JAMA Intern Med* 2021; 181 (7): 1000-1.
30. Cliff AK, Coupland CAC, Keogh RH, et al. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study. *BMJ* 2020; 371: m3731.
31. Nafilyan V, Humberstone B, Mehta N, et al. An external validation of the QCovid risk prediction algorithm for risk of mortality from COVID-19 in adults: national validation cohort study in England. *medRxiv* 2021: 2021.01.22.21249968.
32. Rodriguez-Nava G, Yanez-Bello MA, Trelles-Garcia DP, Chung CW, Friedman HJ, Hines DW. Performance of the quick COVID-19 severity index and the Brescia-COVID respiratory severity scale in hospitalized patients with COVID-19 in a community hospital setting. *Int J Infect Dis* 2021; 102: 571-6.
33. Jr JTK, Yoon JS, Rentsch CT, et al. Development and validation of a 30-day mortality index based on pre-existing medical administrative data from 13,323 COVID-19 patients: the Veterans Health Administration COVID-19 (VACO) Index. *PLoS One* 2020; 15 (11): e0241825.
34. Joseph T King J, Yoon JS, Bredl ZM, et al. Accuracy of the Veterans Health Administration COVID-19 (VACO) Index for predicting short-term mortality among 1307 US academic medical centre inpatients and 427 224 US Medicare patients. *J Epidemiol Community Heal* 2021; 0: jech-2021-216697.
35. Yan L, Zhang H-T, Goncalves J, et al. An interpretable mortality prediction model for COVID-19 patients. *Nat Mach Intell* 2020 25 2020; 2 (5): 283-8.
36. Barish M, Bolourani S, Lau LF, Shah S, Zanos TP. External validation demonstrates limited clinical utility of the interpretable mortality prediction model for patients with COVID-19. *Nat Mach Intell* 2020 31 2020; 3 (1): 25-7.
37. Jimenez-Solem E, Petersen TS, Hansen C, et al. Developing and validating COVID-19 adverse outcome risk prediction models from a bi-national European cohort of 5594 patients. *Sci Reports* 2021 111 2021; 11 (1): 1-12.
38. An C, Lim H, Kim D-W, Chang JH, Choi YJ, Kim SW. Machine learning prediction for mortality of patients diagnosed with COVID-19: a nationwide Korean cohort study. *Sci Reports* 2020 101 2020; 10 (1): 1-11.
39. Gong K, Wu D, Arru CD, et al. A multi-center study of COVID-19 patient prognosis using deep learning-based CT image analysis and electronic health records. *Eur J Radiol* 2021; 139: 109583.
40. Zhang K, Liu X, Shen J, Li et al. Clinically applicable AI System for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. *Cell* 2020; 181 (6): 1423-33.e11.

Capsule

Effect of colonoscopy screening on risks of colorectal cancer and related death

Bretthauer and colleagues performed a pragmatic, randomized trial involving presumptively healthy men and women 55–64 years of age drawn from population registries in Poland, Norway, Sweden, and the Netherlands between 2009 and 2014. The participants were randomly assigned in a 1:2 ratio to either receive an invitation to undergo a single screening colonoscopy (the invited group) or to receive no invitation or screening (the usual-care group). The primary end points were the risks of colorectal cancer and related death, and the secondary end point was death from any cause. Follow-up data were available for 84,585 participants in Poland, Norway, and Sweden: 28,220 in the invited group (11,843 of whom (42.0%) underwent screening) and 56,365 in the usual-care group. A total of 15 participants had major bleeding after polyp removal. No perforations or screening-related deaths occurred within 30

days after colonoscopy. During a median follow-up of 10 years, 259 cases of colorectal cancer were diagnosed in the invited group compared with 622 cases in the usual-care group. In intention-to-screen analyses, the risk of colorectal cancer at 10 years was 0.98% in the invited group and 1.20% in the usual-care group, a risk reduction of 18% (risk ratio [RR] 0.82, 95% confidence interval [95%CI] 0.70–0.93). The risk of death from colorectal cancer was 0.28% in the invited group and 0.31% in the usual-care group (RR 0.90, 95%CI 0.64–1.16). The number needed to invite to undergo screening to prevent one case of colorectal cancer was 455 (95%CI 270–1429). The risk of death from any cause was 11.03% in the invited group and 11.04% in the usual-care group (RR 0.99, 95%CI 0.96–1.04).

N Engl J Med 2022, doi: 10.1056/NEJMoa2208375

Eitan Israeli

Capsule

Of myocarditis and women

Inflammation of the heart muscle, or myocarditis, is an off-target effect of anticancer treatments such as immune checkpoint inhibitor (ICI) therapy. Studying sex differences in ICI myocarditis, **Zhang** et al. observed greater T cell infiltration and cardiac dysfunction in female mice with this disorder and identified down-regulation of two genes, *MANF* and *HSPA5*, in the heart. In a mouse model, cardiac depletion of *Manf* worsened ICI myocarditis, whereas addition of the recombinant *MANF* protein improved heart function.

Treating cardiomyocytes derived from human induced pluripotent stem cells with estrogen or estrogen receptor β agonist induced *MANF* and *HSPA5* expression. Treating female tumor-bearing mice with estrogen receptor β agonist during ICI treatment reduced T cell infiltration and preserved heart function. Hormone therapy could thus potentially limit ICI myocarditis by promoting the protective effects of *MANF*.

Sci Transl Med 2022; 14: abo1981

Eitan Israeli