Immune-Based Prediction of COVID-19 Severity and Chronicity Decoded Using Machine Learning

Bruce K Patterson[1], Jose Guevara-Coto[3], Ram Yogendra, Edgar Francisco, Emily Long, Amruta Pise, Hallison Rodrigues, Purvi Parikh, Javier Mora[2], Rodrigo A Mora-Rodríguez[2]

[1]IncellDx Inc, San Carlos, CA

[2]Lab of Tumor Chemosensitivity, CIET / DC Lab, Faculty of Microbiology, Universidad de Costa Rica

[3]Department of Computer Science and Informatics (ECCI), Universidad de Costa Rica

Summary: Immunologic Modeling of Severity and Chronicity of COVID-19

Corresponding author:
Bruce K. Patterson MD
1541 Industrial Road
San Carlos, CA 94070
Tel: +1.650.777.7630
Fax: +1.650.587.1528
Email: brucep@incelldx.com

Key words:

COVID-19, long haulers, chronic COVID, immune profile, cytokines, chemokines

Abbreviations: IL-interleukin; RANTES-regulation on activation, normal T-expressed and secreted; CCR-chemokine receptor; IFN-interferon, TNF-tumor necrosis factor; MIP-macrophage inflammatory protein; GM-CSF-granulocyte-macrophage colony-stimulating factor; VEGF-vascular endothelial growth factor; HIV; human immunodeficiency virus; HCV hepatitis C virus

ABSTRACT

Individuals with systemic symptoms long after COVID-19 has cleared represent approximately ~10% of all COVID-19 infected individuals. Here we present a bioinformatics approach to predict and model the phases of COVID so that effective treatment strategies can be devised and monitored. We investigated 144 individuals including normal individuals and patients spanning the COVID-19 disease continuum. We collected plasma and isolated PBMCs from 29 normal individuals, 26 individuals with mild-moderate COVID-19, 25 individuals with severe COVID-19, and 64 individuals with Chronic COVID-19 symptoms. Immune subset profiling and a 14-plex cytokine panel were run on all patients. Data was analyzed using machine learning methods to predict and distinguish the groups from each other.Using a multi-class deep neural network classifier to better fit our prediction model, we recapitulated a 100% precision, 100% recall and F1 score of 1 on the test set. Moreover, a first score specific for the chronic COVID-19 patients was defined as **S1 = (IFN-γ + IL-2)/ CCL4-MIP-1β**. Second, a score specific for the severe COVID-19 patients was defined as **S2 = (10*IL-10 + IL-6) – (IL-2 + IL-8)**. Severe cases are characterized by excessive inflammation and dysregulated T cell activation, recruitment, and counteracting activities. While chronic patients are characterized by a profile able to induce the activation of effector T cells with pro-inflammatory properties and the capacity of generating an effective immune response to eliminate the virus but without the proper recruitment signals to attract activated T cells.

INTRODUCTION

Chronic COVID-19 is a group of previously infected individuals, so called "Long Haulers", who experience a multitude of symptoms from several weeks to months after recovering from their acute illness and presumably months after viral clearance. These symptoms include joint pain, muscle aches, fatigue, "brain fog" and others. These symptoms can commonly resemble rheumatic diseases such as rheumatoid arthritis, autoimmune disorders, and others such as fibromyalgia and chronic fatigue syndrome (1). Many of these common disorders are caused by inflammation, hyper- and/or auto-immunity and some such as chronic fatigue are associated with viral persistence after an acute infection with pathogens such as Epstein Barr and Cytomegalovirus (2). Recent studies including those from our laboratory have suggested that (CC) may be caused by persistent COVID itself (3). Here, we sought to identify possible immunologic signatures of COVID-19 severity and to determine whether Chronic COVID-19 might represent a distinct immunologic entity compared to mild to moderate (MM) or severe/critical COVID-19. Further, we addressed the question whether the immunologic profile represents an immune response indicative of prolonged or chronic antigenic exposure. Using machine learning, we identified algorithms that allowed for accurate determination of chronic COVID and severe COVID immunotypes. Further, we present a quantitative immunologic score that could be used to stratify patients to therapy and/or non-subjectively measure response to therapy.

114   **RESULTS**

115   *Immune Profiling*

116   To determine if immunologic abnormalities remain in Long Haulers, we performed high

117   parameter immune cell quantification and characterization in a subset of individuals with

118   preserved peripheral blood mononuclear cells. We determined B-cells, T-cells, and

119   monocytes including subsets and including CD4/CD8 activation and exhaustion. Unlike

120   active COVID-19, the CD4 and CD8 T-cell populations were within normal limits and

121   there was no evidence of T-cell exhaustion (co-expression of PD-1, LAG3, and or

122   CTLA-4). B-cells were significantly elevated compared to normal individuals (P<0.001)

123   as was the CD14+, CD16+ monocytic subset (P<0.001) (Table 1). Interestingly, these

124   two immune cell populations have been shown to be chronically infected by different

125   viruses. B-cells are infected by Epstein-Barr and the CD14+, CD16+ monocytic subset

126   by HIV-1 and by HCV (4).

127   To further characterize the immune response in Long Haulers, we performed

128   quantitative, multiplex cytokine/chemokine panel on 30 normal individuals to establish

129   the normal range of the assay. We then analyzed 64 long haulers and compared the

130   cytokine/chemokine profile (Table 1). IL-2, IL-4, CCL3, IL-6, IL-10, IFN-γ, and VEGF

131   were all significantly elevated compared to normal control (all P<0.001). Conversely

132   GM-CSF and CCL4 were significantly lower than normal controls. Further exacerbating

133   this hyper-immunity was the significant decrease in T regulatory cells compared to

134   normal individuals (P<0.001).

135

136   *Random Forest Binary and Multi-Class Models for Feature Selection and Prediction*

137   We separated the dataset into a training and test split of 90% training and 10% test.

138   This proportion was used because of the reduced number of instances in the dataset.

139   Also, to ensure reproducible results we set the same random seed for all the models.

140   The first model we constructed was the multi-class predictor. This model attempted to

141   separate the severe, long hauler and non-severe-non-long hauler class. This classifier

142   achieved 97% precision, 97% recall and a F1 score of 0.97 in the training partition. In

143   the test split, it performed slightly better, with a precision of 100%, a recall of 100% and

144   thus and F1 score of 1.00 (Table 2). This model was then analyzed to identify the most

145   relevant or informative features. This resulted in the identification of 6 features with an

146   importance score above the importance median (0.063895) and average (0.07143). The

147   identified features were: IFN-γ, IL-2, IL-6, IL-10, IL-8, CCL4-MIP-1β, in importance

148   order. The full list of ranked features can be seen in figure 2.

149   Regarding the long hauler and non-long hauler binary classifier, our results were

150   consistent between the training and the test set. In both partitions the precision and the

151   recall were 100% (1.00) and thus the F1 score equaled 1.00. The observation that the

152   model had good metrics in the test split when compared to the train set is a valuable

153   indicator that the model is not overfitting, and that it is capable of generalizing the

154   patters identified in the training data. The overview of the precision, recall and F1 score

155   for the binary long hauler model can be seen in table 2. Feature importance analysis of

156   the binary model, revealed that the features identified as important for this model were

157   the same features identified as important for the multi-class predictor. This finding

158   suggests there is an important group of characteristics or variables that are influential in

159   the identification of long hauler data points from other instances. These features can be

160   seen in figure 2.

161   The severe binary model, which classified instances between non-severe and severe

162   resulted in high performance metrics for both the training and test splits. As shown in

163   table 2, the performance of this model was an indicator of no potential overfitting. This

164   model is of special interest given the small number of instances in the severe class.

165   Furthermore, the feature importance analysis of this model revealed that the relevant

166   features were also the same as with the multi-class model and with the long hauler

167   binary classifier (Figure 2). This finding reinforces our notion that these group of

168   relevant features could impact classification, or that could have some biological

169   significance worth exploring by means of other analysis like a separation heuristic.

170

171   *Deep Neural Network Binary Classifiers using the Full Feature Set*

172   The deep neural network (DNN) classifier was constructed layers of neurons. Each

173   layer transformed the inputs inputs using the rectified linear activation function or ReLU.

174   The DNN model was constructed to have 1 input layer, 3 hidden layers with 10 neurons

175   each, followed by layer with 6 neurons. Finally, the output layer consists of 3 neuros, for

176   the outputs (classes) and the softmax (multi-class) or sigmoid (binary) function. This

177   architecture was used for the multi-class model and the binary models.

178

179   The results of the long hauler binary models, revealed differences of ~5% between the

180   metrics of the training and the test set (Table 3). Such difference is not significant to

181   attribute overfitting to the training set. In contrast, the severe binary model had

182   significant differences between the performance metrics of the training and the test set

183   (Table 3). This is evident in the precision score, with 98% in the training set and 75% on

184   the test set, and thus the F1 score with a difference of 20% (0.99 on the training set and

185   0.79 on the test set). A potential explanation could be that the severe class has a limited

186   number of data points, but our random forest classifier for the severe class perfumed

187   well. These results suggest that the best approach is a multi-class predictor.

188

189   *Multi-class Deep Neural Network Classifiers using the Full Feature Set*

190   The multi-class DNN implemented using the full feature set had good metrics (Table 3).

191   The precision, recall and F1 score of 100%, 100% and 1.00 in the test split. This

192   indicates that the model is not overfitting, and validating our notion that this would

193   generalize better than the binary models. The model's performance is supported by its

194   confusion matrix (true class vs predicted) where it is possible to determine how well it

195   can predict the three classes (Figure 3).

196   The potential of a DNN classifier is that it adjusts multiple parameters transform the

197   inputs into outputs. This is very important because the vast number of parameters

198   allows for the model to better identify hidden signals in the data. Also, DNN require

199   hyperparameter tuning, such as learning rate, number of hidden layers and neurons per

200   hidden layer, as well as the optimizer and activation function, which affect the

201 performance of the model. By adjusting these hyperparameters and castrating a model

202 capable of finding the hidden relationships in the data we were able to achieve such

203 high results and construct a predictive multi-class system.

204 *Reduced Feature Multi-class Deep Neural Network Classifiers*

205 The results of the DNN indicated that the multi-class had the highest performance.

206 Based on this, we constructed a DNN using the 6 most important features identified by

207 the random forest variable importance. This model was known as minimal DNN or

208 mDNN. This model was constructed using the same architecture as the full feature set

209 DNN. This model's performance in the training set and the test set (Table 4), revealed a

210 significant difference in both precision and recall, such difference could indicate that

211 although the 6 features were identified as the most relevant, it could be possible that all

212 variables contribute to the hidden pattern that makes up the classification of the

213 instances. This idea is supported by the differences in performance between the mDNN

214 and the full feature classifier in both training and test splits (Tables 3 & 4). This is further

215 supported by the comparison of the confusion matrices, where mDNN (figure 4A)

216 misclassifies more instances than the full feature multi-class DNN (Figure 3).

217 Moreover, we simplified our prediction model by feature engineering of two classification

218 scores based on the top informative features. First, a "Long Hauler Score" was defined

219 as S1 = (IFN-γ + IL-2) / CCL4-MIP-1β. Second, "Severe Score" was defined as S2 =

220 (10*IL-10 + IL-6) − (IL-2 + IL8). Using a combined heuristic to first classify the Long

221 Haulers (S1>0.4) and second the severe COVID-19 patients (S2>0), we obtained a

222   sensitivity of 97% for Long Haulers with a 100% specificity and a sensitivity of 88% for

223   severe patients with a specificity of 96% (Figure 4B).

224

225   **DISCUSSION**

226

227   Individuals infected with SARS-Cov2 exert distinct severity patterns which have been

228   associated with different immune activation profiles. Interestingly, in some cases longer

229   times are required to experience full recovery, representing a particular pathological

230   type recently described as long-COVID or long haulers (LH). The scientific evidence

231   generated during the last months strongly supports that the different outcomes on

232   COVID-19 patients are determined by the immune mechanisms activated in response to

233   the viral infection.

234   The immune response to SARS-Cov2 induces a release of different molecules with

235   inflammatory properties such as cytokines and chemokines. This event, known as

236   cytokine storm, is an immunopathological feature of COVID-19 and it has been

237   associated with the severity of the disease. The increase in blood concentrations of

238   different cytokines and chemokines such as IL-6, IL-8, IL-10, TNF-α, IL-1β, IL-2, IP-10,

239   MCP-1, CCL3, CCL4, and CCL5 has been described for COVID-19 patients (5). Some

240   of these molecules have been proposed as biomarkers to monitor the clinical evolution

241   and to determine treatment selection for COVID-19 patients. Nevertheless, it is

242   important to consider that some of these molecules function in a context dependent

243   manner, therefore the clinical relevance of analyzing single cytokine changes is limited.

244   One of the most important challenges during the pandemics is to avoid the saturation of

245   the health systems, therefore the determination of predictive biomarkers that allow a

246   better stratification of the patients is paramount. Even though cytokines such as IL-6

247   and IL-8 have been proposed as indicators of the disease severity, and in some studies

248   they were strong and independent predictors of patient survival (6), their predictive

249   value when analyzed alone is debatable (7). The generation of scores considering blood

250   levels of cytokines and chemokines with different immunological functions incorporates

251   the importance of the context-dependent function of these molecules.

252   In order to predict severe cases, a score was generated considering IL-10, IL-6, IL-2,

253   and IL-8 blood concentrations. In this classification, severe cases are characterized by

254   high IL-6 and IL-10 levels, both cytokines previously attributed to increase the

255   immunopathogenesis of COVID-19 and predictive value in severe cases (6, 8). In

256   different settings, IL-6 has been associated with oxidative stress, inflammation,

257   endothelial dysfunction, and thrombogenesis (9-12) which are characteristic features of

258   severe COVID-19 cases caused by excessive myeloid cell activation (13). Consistently,

259   increased IL-10 levels interfere with appropriate T-cell responses, inducing T-cell

260   exhaustion and regulatory T cell polarization leading to an evasion of the antiviral

261   immune response (14). Furthermore, besides its anti-inflammatory function on T cells, in

262   some settings IL-10 induces STAT1 activation and a pro-inflammatory response in type

263   I IFN-primed myeloid cells (15,16). Therefore, elevated levels of IL-6 and IL-10 promote

264   myeloid cell activation, oxidative stress, endothelial damage, and dampens adequate T

265   cell activation. Additionally, to strengthen the classification, the score presented here,

266   differentiates the severe cases by the subtraction of IL-2 and IL-8, which are cytokines

267   related to proper T cell activation (IL-2) and recruitment (IL-8).

268   According to the score generated for distinguishing LH, these patients are characterized

269   by an increased IFN-γ and IL-2 and a reduced CCL4 production. In the context of a viral

270   infection, the combination of IFN-γ and IL-2 would induce the activation of effector T

271   cells with pro-inflammatory properties and the capacity of generating an effective

272   immune response to eliminate the virus. However, LH are characterized by longer

273   periods of time with clinical signs and symptoms such as fatigue and lung damage. This

274   suggests that the inflammatory context created by these cytokines to induce T cell

275   activation is not enough to generate an adequate anti-viral response without the proper

276   recruitment signals to attract activated T cells. CCL4 signals through the receptor CCR5

277   to attract T cells to the site of inflammation and depending on the immune context, this

278   molecule recruits differently activated T cells (17,18). Moreover, it was recently shown

279   by single cell analysis a down regulation of CCL4 expression in peripheral myeloid cell

280   compartments in patients with mild and severe COVID-19 (19). In LH, IFN-γ and IL-2

281   would create an immune context to induce Th1 polarization, but the low levels of CCL4

282   affect the recruitment of these cells impairing the antiviral response. The effect of

283   increased IFN-γ and IL-2 on T cell activation is evident in the reduction of the

284   percentage of exhausted (CD4+PD1+/ CD8+PD1+) and regulatory T cells (FoxP3+)

285   compared to healthy donors. Interestingly, there is an increase in the percentage of

286   circulating CD4+ and CD8+ T cells expressing CTLA-4 in the LH group compared to

287   healthy donors, which is a molecule that affects antigen presentation in secondary

288   lymphoid organs, but its presence in circulating T cells may reflect a compensatory

289   mechanisms to the low CCL4 levels in the LH group. CTLA-4 induced signaling in T

290   cells upregulates the expression of the CCL4 receptor CCR5 (20, 21), in the LH group

291   CTLA-4 upregulation suggests a failed attempt to increase the sensitivity of IFN-γ/IL-2

292   activated T cells to CCL4. Therefore, proper T cell activation (high IFN-γ+IL-2) but

293   ineffective T cell recruitment (low CCL4) are characteristic features of the failed anti-

294   viral response observed in the LH group supporting virus persistence. Additionally,

295   increased IFN-γ promotes myeloid cell activation which is observed in the augmented

296   percentage of inflammatory CD14+CD16+ monocytes in the LH group compared to

297   healthy donors, supporting lymphopenia and virus persistence in these patients. This is

298   supported by recent findings describing an increased gene expression in response to

299   IFN-γ in mild and severe COVID-19 patients in peripheral myeloid cells (19) and the

300   dysregulation in the balance of monocyte populations by the expansion of the monocyte

301   subsets described in COVID-19 patients (22). Finally, we propose that long-lasting

302   pulmonary damage observed in LH, is caused by a combination of factors including 1)

303   longer virus persistence influenced by LH immune profile characterized by high IFN-γ

304   and IL-2 levels inducing Th1 polarization which is ineffective with low CCL4-induced T

305   cell recruitment, leading to an inflammatory myeloid cell activation; and 2) the

306   immunopathological pulmonary effects consequence of this LH immune profile.

307   Regarding the immunopathological effects of LH immune profile, using murine models it

308   has been shown that high IFN-γ levels could affect the kinetics of the resolution of

309   inflammation-induced lung injury as well as thrombus resolution (23, 24), which could be

310   related to long-lasting symptoms of LH associated to pulmonary coagulopathy and

311   immune-mediated tissue damage.

312 Interestingly, COVID-19 individuals (including LH, mild, severe) show high levels of

313 CCL5, a chemoattractant that like CCL4 signals through CCR5. Indeed, the disruption

314 of the CCL5-CCR5 pathway restores immune balance in critical COVID-19 patients (4).

315 In the specific case of LH, despite the high concentrations of CCL5 a reduction on the

316 CCL4-mediated recruitment of activated T cells is proposed. This could be related to

317 different factors:

318 (1) Reduction of total recruitment signals in LH with low CCL4 concentrations.

319 (2) Different functional responses of CCL4 and CCL5 to polymorphic variants of the

320 CCR5. Distinct functional features have been reported to CCR5 variants regarding

321 binding avidity, receptor internalization, Ca++ influx and chemotactic activity (25). Even

322 though, clear mechanistic differences between CCL4 and CCL5 interaction with CCR5

323 are missing, it has been suggested that is important to consider the knowledge gained

324 on CCR5 polymorphisms in HIV/AIDS context (26).

325 (3) Signaling through alternative receptors for CCL5. Besides CCR5, CCL5 can signal

326 through the receptors CCR1 and CCR3 (27) whereas CCL4 effects are restricted to

327 CCL5. It has been shown that CCL4 can bind to CCR1 but is not able to induce the

328 intracellular pathway necessary for activating the chemoattractant stimulus (27,28) .

329 Therefore, CCL4 has been proposed as an antagonist of CCR1 (28), however further

330 analysis of this needs to be performed. Interestingly, CCR1 is expressed on blood

331 myeloid cells such as monocytes and neutrophils (27), and it is upregulated on COVID-

332 19 patients (29). Additionally, high levels of IFN-γ (feature of LH) have been associated

333 with an increase CCR1 expression on human neutrophils (30). Therefore, in LH, high

334   levels of CCL5 (combined with low levels of potential CCR1-antagonist CCL4) leads to

335   a higher recruitment of myeloid cells expressing CCR1.

336

337   **MATERIAL/METHODS**

338   *Patients*

339   Following informed consent, whole blood was collected in a 10 mL EDTA tube and a 10

340   mL plasma preparation tube (PPT). A total of 144 individuals were enrolled in the study

341   consisting of 29 normal individuals, 26 mild-moderate COVID-19 patients, 25 severe

342   COVID-19 patients and 64 chronic COVID (long hauler-LH) individuals. Long Haulers

343   symptoms are listed in Figure 1. Study subjects were stratified according to the

344   following criteria.

345   Mild

346   1. Fever, cough, sore throat, malaise, headache, myalgia, nausea, diarrhea, loss of

347   taste and small

348   2. No sign of pneumonia on chest imaging (CXR or CT Chest)

349   3. No shortness of breath or dyspnea

350   Moderate:

351   1. Radiological findings of pneumonia fever and respiratory symptoms

352   2. Saturation of oxygen (SpO2) ≥ 94% on room air at sea level

353   Severe

354   1. Saturation of oxygen (SpO2) < 94% on room air at sea level

355   2. Arterial partial pressure of oxygen (PaO2)/ fraction of inspired oxygen (FiO2) <

356   300mmHG

357   3. Lung infiltrate > 50% within 24 to 48 hours

358   4. HR ≥ 125 bpm

359   5. Respiratory rate ≥ 30 breaths per minute

360 <u>Critical</u>

361   1. Respiratory failure and requiring mechanical ventilation, ECMO, high-flow nasal

362   cannula oxygen supplementation, noninvasive positive pressure ventilation

363   (BiPAP, CPAP)

364   2. Septic Shock- Systolic blood pressure < 90mmHg or Diastolic blood pressure <

365   60 mmHg or requiring vasopressors (levophed, vasopressin, epinephrine

366   3. Multiple organ dysfunction (cardiac, hepatic, renal, CNS, thrombotic disease)

367

368 <u>Post-acute COVID-19 (Long COVID)</u>

369   1.   Extending beyond 3 weeks from the initial onset of first symptoms

370 <u>Chronic COVID-19</u>

371   1. Extending beyond 12 weeks from the initial onset of first symptoms (Table 1)

372

373 *High Parameter Immune Profiling/Flow Cytometry*

374 Peripheral blood mononuclear cells were isolated from peripheral blood using

375 Lymphoprep density gradient (STEMCELL Technologies, Vancouver, Canada). Aliquots

376 200 of cells were frozen in media that contained 90% fetal bovine serum (HyClone,

377 Logan, UT) and 10% dimethyl sulfoxide (Sigma-Aldrich, St. Louis, MO) and stored at -

378 70°C. Cells were stained and analyzed as previously described (4) (Patterson) using a

379 17-color antibody cocktail.

380

381 *Multiplex Cytokine Quantification*

382 Fresh plasma was used for cytokine quantification using a customized 14-plex bead

383 based flow cytometric assay (IncellKINE, IncellDx, Inc) on a CytoFlex flow cytometer as

384 previously described using the following analytes: 'TNF-α', 'IL-4', 'IL-13','IL-2', 'GM-

385 CSF', 'sCD40L', 'CCL5 (RANTES)', 'CCL3 (MIP-1α)','IL-6', 'IL-10', 'IFN-γ', 'VEGF', 'IL-

386 8', and 'CCL4 (MIP-1β) (4). For each patient sample, 25 μL of plasma was used in each

387 well of a 96-well plate. Standard curves with serial 6 point dilutions of antigen were run

388 on each plate for each cytokine. Raw data was analyzed using LegendPlex software

389 (Biolegend, Inc San Diego CA). Samples were run in duplicate.

390

391 *Data Processing*

392 Data was imported and processed using Python 2.7, using the *pandas* library (version

393 1.1.0). and the numeric python module, *numpy* version 1.18.5. Our data consisted of

394 144 instances representing 4 classes (Normal-n=29, Mild-Moderate-n=26, Severe-n=25,

395 Long Hauler-n=64). Each class had 14 columns, representing the different

396 cytokine/chemokine analytes. Each analyte had different measurements which required

397 a normalization process to reduce outlier effect and to facilitate algorithm convergence.

398   Normalization was done using Min-Max and based on a linear transformation of the

399   original data. Min-Max maintains the original relationship between the data, while fitting

400   it within a pre-defined boundary. The Python implementation of min-max calculates the

401   range in such a manner that the range of the features will be defined between 0 and 1.

402   For this reason, min-max normalization is also referred to as 0-1 normalization (or

403   scaling). The typical min-max transformation is given in equation 1:

404

405    $$X = \frac{(X - Xmin)}{Xmax - Xmin} \qquad [1]$$

406

407

408   *Target Variable Processing*

409   Since Min-max normalization, can only be applied to numeric variables a new variable

410   defined as *targets* was created.The variable targets represent the different classes

411   (Long Hauler, Severe, Mild-Moderate, and Normal) for the instances in the dataset. The

412   resulting array has 4 classes for each state. The goal of our analysis is to properly

413   identify/discriminate the instances that belong to the Severe state or the Long-Hauler

414   state compared to other states. This goal can be achieved by building either binary

415   classifiers for the Severe class and for the Long Hauler class, a multi-class predictor.

416   For the construction of both models, t is required to separate the targets to reflect the

417   dosing question: can a predictor discriminate between the Severe, Long Hauler and

418   Other Sates.

419 To build the models that answer this question, we grouped the M-M and Normal labels

420 in a new class which was distinct form the Severe and Long-Hauler states. We then

421 proceeded to apply filters based on the task (binary or multi-class classification). For the

422 Severe binary predictor, we conditioned the targets to be exactly Severe or else they

423 were assigned to Not-Severe. This same task was done for Long-Haulers, were either

424 an instance label was exactly labelled Long-Hauler or else it would be assigned to the

425 Non-Long Hauler class. The multi-class predictor processing only requires to define

426 three classes: Severe, Long-Hauler and Non-Severe-Non-Long-Hauler which was

427 composed of the Normal and Mild-Moderate cases.

428

429 *One-hot Encoding of Targets*

430 The implementation of one-hot encoding on the target variable, is based on the notion

431 that multiple machine learning algorithms are unable to properly process categorical

432 data. It is possible to use numeric replacements, such as integer values, but this can

433 only be useful if there is an ordinal relationship within the variable. Such use would

434 imply that there exists a vectorial relationship between the labels, for example, in our

435 classes we have Normal, Mild-Moderate, Severe and Long-Haulers. If we assigned a

436 vector of integers from 0 to 4 in their corresponding orders to the classes, it would

437 assume the presence of a vectorial distance between Normal and Long Hauler or V0 ->

438 V4.

439 To properly design an experiment that reflects this, we use one-hot encoding After

440 applying one-hot encoding the labels are substituted with 1 and 0, where 1 represents

441   the presence of the class and 0 the absence. The use of one-jot encoding corrects for

442   the vector-distance assumption of integer or categorical classes, where higher or larger

443   values could be interpreted as better.

444

445   Definition of precision, recall and F1 score

446   The precision (equation 2) is a measure of the percentage of the results that are

447   relevant. The metric Recall measures the percentage of the total relevant results that

448   are correctly classified by the predictor (equation 3). The harmonic mean between these

449   two measures is known as the F1 score and ranges from 0 to 1, the closer to is to 1, the

450   better the model performs (equation 4). The F1 score for both false positives (FP) and

451   false negatives (FN) as well as for true positives (TP).

452
$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \qquad [2]$$

453
$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \qquad [3]$$

454
$$F1 = \frac{2*Precision*Recall}{Precision + Recall} = \frac{TP}{TP + 1/2(FP+FN)} \qquad [4]$$

455

456   *Feature Selection and Classification using Random Forest*

457   Data pre-processing, target variable processing and the encoding of targets were

458   performed before classification as above. Feature selection is the process of reducing

459   dimensionality of the dataset by selecting those features or variables that are more

460   informative than those that are not.

461   To perform feature selection, we implemented the RandomForestClassifier method from

462   Sci-kit Learn. Random Forest allows for identification of features that better separate the

463   classes by determining what percentage of the nodes that use those features have a

464   reduction in entropy or impurity (which are measures of how well separated the

465   instances are using a feature).

466   The binary classifier was constructed using the data points and their features with the

467   one-hot encoded target corresponding to: 1) the severe and non-severe model, 2) the

468   long hauler and non-long hauler model and 3) the multiclass model. The model was

469   built with the RandomForestClassifier method from Sci-kit Learn, with the number of

470   trees constructed set to 750, the number of features set as the square root of the

471   feature space, and the node depth equal to 4 to avoid overfitting. These parameters

472   were set for binary and multi-class predictors. Model performance was measured using:

473   precision, recall and the F1 score (see supplementary information).

474

475   *Predictor Construction Using Deep Neural Networks*

476   The deep neural network (DNN) binary and multiclass classifiers were constructed with

477   a basic DNN architecture built on stacks of perceptrons, where each subsequent layer

478   is connected to the previous one. Each layer transformed the inputs inputs using the

479   rectified linear activation function or ReLU. The DNN models were constructed to have

480   1 input layer, 3 hidden layers with 10 neurons each, followed by layer with 6 neurons.

481   Finally, the output layer consists of 3 neurons, for the outputs (classes) and the softmax

482   (multi-class) or sigmoid (binary) function.

483   In order for a DNN to generate the best possible predictions, we minimized the loss

484   function or error of the model using the ADAM optimizer to search for the optimal

485   combination of hyperparameters. When setting the optimizer, we defined the learning

486   rate to 1e-3. The loss function was set to categorical cross entropy because the targets

487   are one-hot encoded.

488

489

490   **REFERENCES**

491

492   1. L. Chen, H. Deng, H. Cui, J. Fang, Z. Zuo, J. Deng, Y. Li, X. Wang, L. Zhao
493   Inflammatory responses and inflammation-associated diseases in organs. _Oncotarget_
494   **9**, 7204–7218 (2018).

495   2. S. Rasa, Z. Nora-Krukle, N. Henning et al. Chronic viral infections in myalgic
496   encephalomyelitis/chronic fatigue syndrome (ME/CFS). _J Transl Med_ **16**, 268 (2018).

497   3. P.A. Mudd, J.S. Turner, A. Day, W.B. Alsoussi, Z. Liu, J.A. O'Halloran, R.M. Presti,
498   B.K. Patterson,S.P.J. Whelen, A. Ellebedy. SARS-CoV-2 viral RNA shedding for more
499   than 87 days in an individual with an impaired CD8+ T-cell response. _Front Immunol_ (in
500   press).

501   4. G. Coquillard, B. Patterson. HCV-Infected, Monocyte Lineage Reservoirs Differ in
502   Individuals with or without HIV Co-Infection. _J Infect Dis_ **2009**;200:947-954.
503
504   5. B. K. Patterson, H. Seethamraju, K. Dhody, M. J. Corley, K. Kazempour, J. P.,
505   Lalezari, A. P. Pang, C. Sugai, E. B. Francisco, A. Pise, H. Rodrigues, M. Ryou, H. L.
506   Wu, G. M. Webb, B. S. Park, S. Kelly, N. Pourhassan, A. Lelic, L. Kdouh, M. Herrera, E.
507   Hall, E. Aklin, L. Ndhlovu, J. B. Sacha. CCR5 inhibition in Critical COVID-19 Patients
508   Decreases Inflammatory Cytokines, Increases CD8 T-Cells, and Decreases SARS-
509   CoV2 RNA in Plasma by Day 14. Int J Infect Dis (2020) doi: 10.1016/j.ijid.2020.10.101

6. D. M. Del Valle, S. Kim-Schulze, H. H. Huang, N. D. Beckmann, S. Nirenberg, B. Wang, Y. Lavin, T. H. Swartz, D. Madduri, A. Stock, T. U. Marron, H. Xie, M. Patel, K. Tuballes, O. Van Oekelen, A. Rahman, P. Kovatch, J. A. Aberg, E. Schadt, S. Jagannath, M. Mazumdar, A. W. Charney, A. Firpo-Bncourt, D. R. Mendu, J. Jhang, D. Reich, K. Sigel, C. Cordon-Cardo, M. Feldmann, S. Parekh, M. Merad, S. Gnjatic,. An inflammatory cytokine signature predicts COVID-19 severity and survival. *Nat Med*, **26**, 1636-1643 (2020).

7. S. M. Russell, A. Alba-Patiño, E. Barón, M. Borges, M. Gonzalez-Freire, & R. de la Rica. Biosensors for Managing the COVID-19 Cytokine Storm: Challenges Ahead. *ACS Sens*, **5**, 1506-1513.

8. S. K.Dhar, K, V., S. Damodar, S. Gujar & M. Das, IL-6 and IL-10 as redictors of disease severity in COVID 19 patients: Results from Meta-analysis and Regression. *medRxiv*, (2020). 2008.2015.20175844. https://doi.org/10.1101/2020.08.15.20175844

9. T. Hou, Tieu, B. C., S. Ray, A. Recinos Iii, R. Cui, R. G. Tilton, & A. R. Brasier. Roles of IL-6-gp130 Signaling in Vascular Inflammation. *Curr Cardiol Rev*, **4**, 179-192. https://doi.org/10.2174/157340308785160570

10. J. Lee, S. Lee, H. Zhang, M. A. Hill, C. Zhang, & Y. Park. Interaction of IL-6 and TNF-α contributes to endothelial dysfunction in type 2 diabetic mouse hearts. *PLoS One*, **12**, e0187189. https://doi.org/10.1371/journal.pone.0187189

11. V. Roldán, F. Marín, A. D. Blann, A. García, P. Marco, F. Sogorb, & G. Y. Lip. Interleukin-6, endothelial activation and thrombogenesis in chronic atrial fibrillation. *Eur Heart J*, **24**, 1373-1380. https://doi.org/10.1016/s0195-668x(03)00239-2

12. S. Wassmann, M. Stumpf, K. Strehlow, A. Schmid, B. Schieffer, M. Böhm, & G. Nickenig. Interleukin-6 induces oxidative stress and endothelial dysfunction by overexpression of the angiotensin II type 1 receptor. *Circ Res*, **94**, 534-541. https://doi.org/10.1161/01.res.0000115557.25127.8d

13. D. McGonagle, K. Sharif, A. O'Regan, & C. Bridgewood. The Role of Cytokines including Interleukin-6 in COVID-19 induced Pneumonia and Macrophage Activation Syndrome-Like Disease. *Autoimmun Rev*, **19**, 102537. https://doi.org/10.1016/j.autrev.2020.102537

14. J. M. Rojas, M. Avia, V. Martín, & N. Sevilla. IL-10: A Multifunctional Cytokine in Viral Infections. *J Immunol Res*, **2017**, 6104054. https://doi.org/10.1155/2017/6104054

15. H. Mühl. Pro-Inflammatory Signaling by IL-10 and IL-22: Bad Habit Stirred Up by Interferons? *Front Immunol*, **4**, 18. https://doi.org/10.3389/fimmu.2013.00018

555   16. M. N. Sharif, I. Tassiulas, Y. Hu, I. Mecklenbräuker, A. Tarakhovsky, & L. B.
556   Ivashkiv. IFN-alpha priming results in a gain of proinflammatory function by IL-10:
557   implications for systemic lupus erythematosus pathogenesis. *J Immunol*, **172**, 6476-
558   6481. https://doi.org/10.4049/jimmunol.172.10.6476

560   17. J. Y. Liu, F. Li, L. P. Wang, X. F. Chen, D. Wang, L. Cao, Y. Ping, S. Zhao, B. Li, S.
561   H. Thorne, B. Zhang, P. Kalinski, & Y. Zhang. CTL- vs Treg lymphocyte-attracting
562   chemokines, CCL4 and CCL20, are strong reciprocal predictive markers for survival of
563   patients with oesophageal squamous cell carcinoma. *Br J Cancer*, **113,** 747-755.
564   https://doi.org/10.1038/bjc.2015.290

566   18. N. Mukaida, S. I. Sasaki, & T. Baba. CCL4 Signaling in the Tumor
567   Microenvironment. *Adv Exp Med Biol*, **1231**, 23-32. https://doi.org/10.1007/978-3-030-
568   36667-4_3

570   19. G. Xu, F. Qi, H. Li, Q. Yang, H. Wang, X. Wang, X. Liu, J. Zhao, X. Liao, Y. Liu, L.
571   Liu, S. Zhang, & Z. Zhang. The differential immune responses to COVID-19 in
572   peripheral and lung revealed by single-cell RNA sequencing. *Cell Discov*, **6**, 73.
573   https://doi.org/10.1038/s41421-020-00225-2

575   20. K. Knieke, H. Hoff, F. Maszyna, P. Kolar, A. Schrage, A. Hamann, G. F. Debes, M.
576   C. Brunner-Weinzierl.. CD152 (CTLA-4) determines CD4 T cell migration in vitro and in
577   vivo. *PLoS One*, **4**, e5702. (2009). https://doi.org/10.1371/journal.pone.0005702

579   21. K. Knieke, H. Lingel, K. Chamaon, & M. C. Brunner-Weinzierl. Migration of Th1
580   lymphocytes is regulated by CD152 (CTLA-4)-mediated signaling via PI3 kinase-
581   dependent Akt activation. *PLoS One*, **7**, e31391. (2012).
582   https://doi.org/10.1371/journal.pone.0031391

584   22. W. Shi, X. Liu, Q. Cao, P. Ma, W. Le, L. Xie, J. Ye, W. Wen, H. Tang, W. Su, Y.
585   Zheng, & Y. Liu. High-dimensional single-cell analysis reveals the immune
586   characteristics of COVID-19. *Am J Physiol Lung Cell Mol Physiol*.
587   https://doi.org/10.1152/ajplung.00355.2020

589   23. J. R. Mock, M. K. Tune, C. F. Dial, J. Torres-Castillo, R. S. Hagan, & C. M.
590   Doerschuk. Effects of IFN-γ on immune cell kinetics during the resolution of acute lung
591   injury. *Physiol Rep*, **8**, e14368. https://doi.org/10.14814/phy2.14368

593   24. M. Nosaka, Y. Ishida, A. Kimura, Y. Kuninaka, M. Inui, N. Mukaida, & T. Kondo.
594   Absence of IFN-γ accelerates thrombus resolution through enhanced MMP-9 and VEGF
595   expression in mice. *J Clin Invest*, **121**, 2911-2920. https://doi.org/10.1172/jci40782

597   25. H. F. Dong, K. Wigmore, M. N. Carrington, M. Dean, J. A. Turpin, & O. M. Howard.
598   Variants of CCR5, which are permissive for HIV-1 infection, show distinct functional
599   responses to CCL3, CCL4 and CCL5. *Genes Immun*, **6**, 609-619.
600   https://doi.org/10.1038/sj.gene.6364247

601 26. R. K. Mehlotra. Chemokine receptor gene polymorphisms and COVID-19: Could
602 knowledge gained from HIV/AIDS be important? *Infect Genet Evol*, *85*, 104512.
603 https://doi.org/10.1016/j.meegid.2020.104512
604
605 27. C. E. Hughes, & R. J. B. Nibbs. A guide to chemokines and their receptors. *Febs j*,
606 *285*, 2944-2971. https://doi.org/10.1111/febs.14466
607
608 28. H. Gaertner, O. Lebeau, I. Borlat, F. Cerini, B. Dufour, G. Kuenzi, A. Melotti, R. J.
609 Fish, R. Offord, J. Y. Springael, M. Parmentier, & O. Hartley. Highly potent HIV
610 inhibition: engineering a key anti-HIV structure from PSC-RANTES into MIP-1 β/CCL4.
611 *Protein Eng Des Sel*, *21*, 65-72. https://doi.org/10.1093/protein/gzm079
612
613 29. P. R. Ray, A. Wangzhou, N. Ghneim, M. S. Yousuf, C. Paige, D. Tavares-Ferreira,
614 J. M. Mwirigi, S. Shiers, I. Sankaranarayanan, A. J. McFarland, S. V. Neerukonda, S.
615 Davidson, G. Dussor, M. D. Burton, & T. J. Price. A pharmacological interactome
616 between COVID-19 patient samples and human sensory neurons reveals potential
617 drivers of neurogenic pulmonary dysfunction. *Brain Behav Immun*, *89*, 559-568.
618 https://doi.org/10.1016/j.bbi.2020.05.078
619
620 30. R. Bonecchi, N. Polentarutti, W. Luini, A. Borsatti, S. Bernasconi, M. Locati, C.
621 Power, A. Proudfoot, T. N. Wells, C. Mackay, A. Mantovani, & S. Sozzani. Up-
622 regulation of CCR1 and CCR3 and induction of chemotaxis to CC chemokines by IFN-γ
623 in human neutrophils. *J Immunol*, *162*, 474-479.

624
625
626
627

632

Author contributions:
634 R.Y. organized the clinical study and actively recruited patients.
635 B.K.P, A.P., H.R., E.L. performed experiments and analyzed the data.
636 J.G-C., R.A.M., J.M. performed the bioinformatics
637 B.K.P., J.M., J.G-C., R.A.M. wrote the draft of the manuscript and all authors
638 contributed to revising the manuscript prior to submission.

639

Competing interests:
641 B.K.P, A.P., H.R., E.L. are employees of IncellDx

642

643 **Data and materials availability**:

644

645 All requests for materials and data should be addressed to the corresponding author

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676 **TABLES and FIGURES**

677

678 Figure 1. Symptoms reported by long hauler patients enrolled in the study.

COVID-19 Long Hauler Symptoms
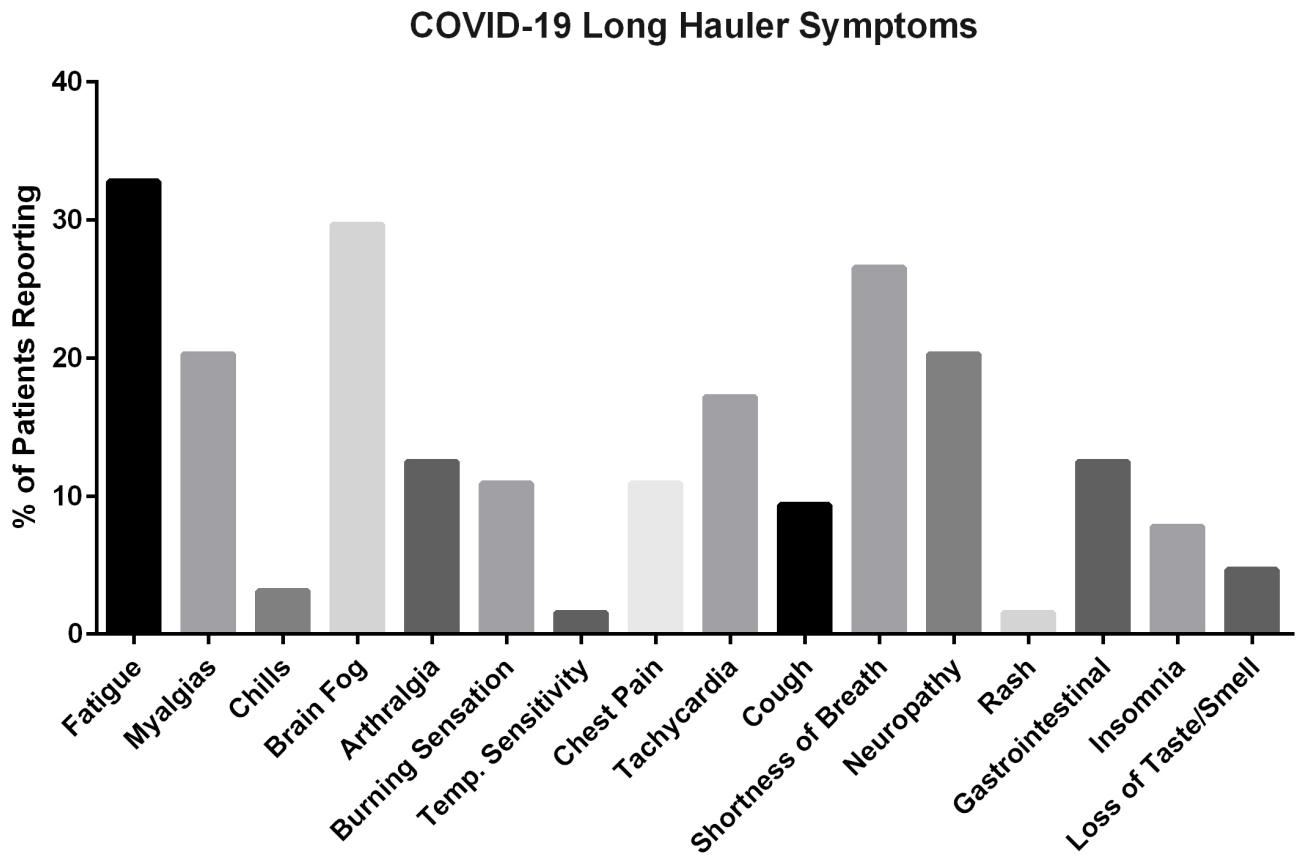
679

680

681

682

683

684

685

686

687

688 **Table 1. Immunologic parameters of study participants**

Average Normals	CD3+%	CD4%	CD8%	CD4+PD1+%	CD4+LAG3+%	CD4+CTLA4+%	CD4+FoxP3+%	CD8+PD1+ %	CD8+LAG3+ %	CD8+CTLA4+%	CD19+%	CD14+CD16-%	CD16+CD14+%	CD16+CD14-%
Average Normals	64.45	53.89	33.83	35.62	0.94	1.51	6.21	43.76	4.35	1.39	6.04	42.79	9.00	32.68
Lower CI	54.39	43.21	27.20	28.36	0.49	0.75	4.54	33.50	2.71	0.74	5.04	34.41	4.60	25.49
Upper CI	74.50	64.57	40.46	42.89	1.39	2.26	7.87	54.01	5.99	2.03	7.04	51.16	13.41	39.86

689

Average Long Hauler	CD3+%	CD4%	CD8%	CD4+PD1+%	CD4+LAG3+%	CD4+CTLA4+%	CD4+FoxP3+%	CD8+PD1+ %	CD8+LAG3+ %	CD8+CTLA4+%	CD19+%	CD14+CD16-%	CD16+CD14+%	CD16+CD14-%
Average Long Hauler	48.98	56.18	35.36	17.78	0.72	4.06	2.58	31.99	0.71	3.11	13.14	19.07	29.30	33.86

690

Average (pg/ml)	TNF-α	IL-4	IL-13	IL-2	GM-CSF	sCD40L	CCL5 (RANTES)	CCL3 (MIP-1α)	IL-6	IL-10	IFN-γ	VEGF	IL-8	CCL4 (MIP-1β)
Normals	9.09	4.18	3.94	6.17	51.27	7192.39	10781.84	22.82	2.21	0.67	1.94	9.32	16.87	76.84
Lower CI	7.37	2.17	1.79	5.53	25.72	5148.85	9764.99	13.05	1.65	0.42	0.63	6.36	13.03	61.00
Upper CI	10.81	6.18	6.09	6.82	76.82	9235.92	11798.68	32.60	2.77	0.92	3.26	12.28	20.72	92.67
Long Haulers	7.72	17.03	4.21	16.16	12.46	18302.41	12505.06	97.81	20.47	12.23	86.60	41.03	35.98	35.10
Mild-Mod	6.82	2.33	2.40	5.90	56.13	10673.72	11627.70	18.75	8.74	0.63	1.15	17.39	17.37	94.40
Severe	5.39	2.39	2.26	5.43	20.31	12306.39	11581.47	16.54	144.15	3.10	2.06	25.52	10.87	64.84

691

692

693

694 **Table 2. Performance Metrics for the Random Forest Classifiers in the test split.**

Model	Precision %	Recall %	F1 Score
Long Hauler-Full Features	100	100	1.00
Severe- Full Features	100	100	1.00
Multi-Class-Full Features	100	100	1.00

695

696   Figure 2. Feature importance for multi-class classifier using Random Forest predictor.
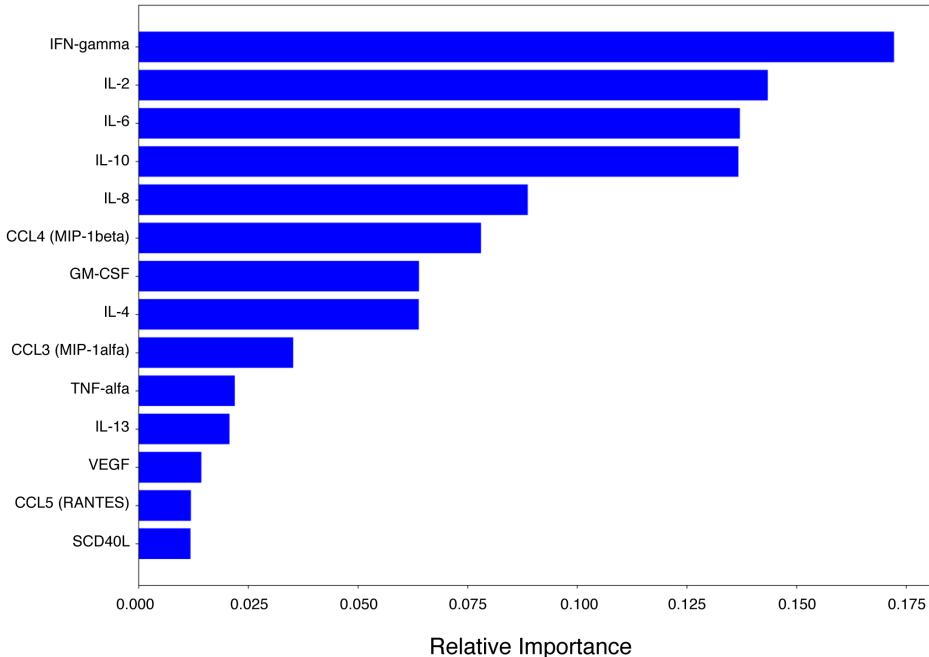
697

698

699

700

701

702

703

704

705   Table 3. Performance Metrics of the DNN full feature model in the training and test splits

706

DNN	Precision %	Recall %	F1 Score
Multi-Class - Full Features -Train	99	97	0.98
Long Hauler - Full Features -Train	100	100	1.00
Severe - Full Features - Train	98	100	0.99
Multi-Class - Full Features -Test	100	100	1.00
Long Hauler - Full Features -Test	94	94	0.93
Severe - Full Features - Test	75	92	0.79

707

708   Figure 3. Full-feature multi-class DNN model confusion matrix for the test split.
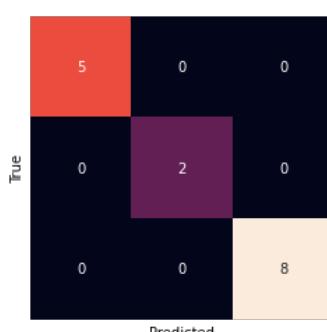

709

710   Table 4. Performance metrics for the minimal deep neural network (mDNN) on the

711   training and test splits.

| Model	Precision %	Recall %	F1 Score
mDNN-Training	98	96	0.97
mDNN- Test	82	89	0.84

712

713
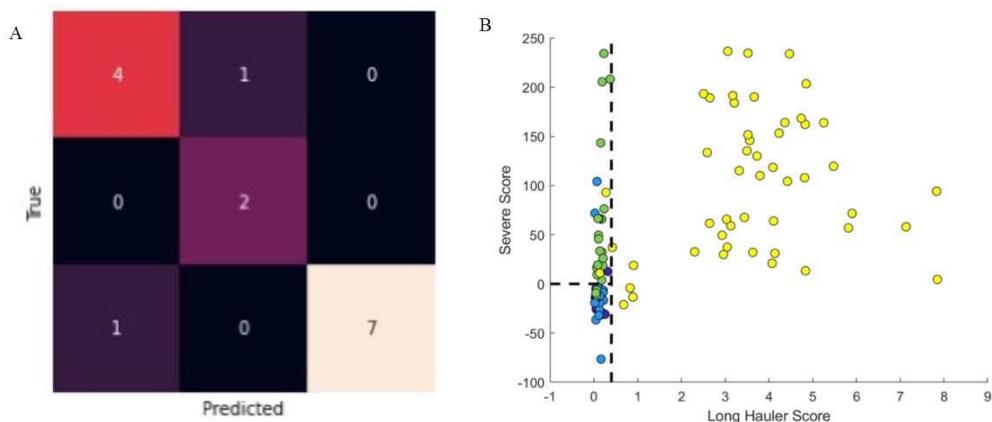
714   Figure 4. Classification abilities of the minimal Deep Neural Network (mDNN) and the

715   discrimination heuristic generated using important variables. A) The confusion matrix for

716   the mDNN classifier denoting the presence of false positives for the severe and other

717   classes. B) Discrimination ability of the heuristic with reduced or most important features

718   identified using Random Forest classifier. The dots represent the data points, where

719   yellow are long haulers, green-severe, dark blue-mild/moderate and light blue-normal.

720