

1 **Machine learning vs. traditional regression analysis for fluid overload prediction in the ICU**

2

3 Andrea Sikora, PharmD, MSCR, BCCCP, FCCM<sup>1</sup>

4 1120 15th Street, HM-118 Augusta, GA 30912

5 (706)-721-3826

6 sikora@uga.edu

7 <sup>1</sup>University of Georgia College of Pharmacy, Department of Clinical and Administrative Pharmacy,

8 Augusta, GA, USA

9

10 Tianyi Zhang<sup>2</sup>

11 Tianyi.zhang@uga.edu

12 <sup>2</sup>University of Georgia Franklin College of Arts and Sciences, Department of Statistics, Athens, GA, USA

13

14 David J. Murphy, MD, PhD<sup>3</sup>

15 david.j.murphy@emory.edu

16 <sup>3</sup>Emory University, Division of Pulmonary, Allergy, Critical Care and Sleep Medicine, Atlanta, GA, USA

17

18 Susan E. Smith, PharmD<sup>1</sup>

19 Susan.smith@uga.edu

20 <sup>1</sup>University of Georgia College of Pharmacy, Department of Clinical and Administrative Pharmacy,

21 Augusta, GA, USA

22

23 Brian Murray, PharmD<sup>4</sup>

24 Brian.murray@unchealth.unc.edu

25 <sup>4</sup>University of North Carolina Medical Center, Department of Pharmacy, Chapel Hill, NC, USA

26

27 Rishikesan Kamaleswaran, PhD<sup>5,6</sup>

28 rkamaleswaran@emory.edu

29 <sup>5</sup>Department of Biomedical Informatics, Emory University School of Medicine, Atlanta, GA,

30 USA

31 <sup>6</sup>Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA, USA

32

33 Xianyan Chen, PhD<sup>2</sup>

34 xychen@uga.edu

35 <sup>2</sup>University of Georgia Franklin College of Arts and Sciences, Department of Statistics, Athens, GA, USA

36

37 Mitchell S. Buckley, PharmD, FCCM<sup>7</sup>

38 mitchellsbuckley@gmail.com

39 <sup>7</sup>LaJolla Pharmaceuticals

40

41 Sandra Rowe, PharmD<sup>8</sup>

42 rowes@ohsu.edu

43 <sup>8</sup>Oregon Health and Science University, Department of Pharmacy, Portland, OR, USA

44

45 John W. Devlin, PharmD, MCCM<sup>\*9,10</sup>

46 j.devlin@northeastern.edu

47 <sup>9</sup>Northeastern University School of Pharmacy, Boston, MA

48 <sup>10</sup>Brigham and Women's Hospital, Division of Pulmonary and Critical Care Medicine, Boston, MA, USA

49

50 On behalf of the MRC-ICU Investigator Team

51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61

**Conflicts of Interest:** The authors have no conflicts of interest.

**Funding:** Funding through Agency of Healthcare Research and Quality for Drs. Devlin, Murphy, Sikora, Smith, and Kamaleswaran was provided through R21HS028485 and R01HS029009.

**Acknowledgements:** Data acquisition were supported by NC TraCS, funded by Grant Number UL1TR002489 from the National Center for Advancing Translations Sciences at the National Institutes of Health, and Data Analytics at the University of North Carolina Medical Center Department of Pharmacy.

62 **ABSTRACT**

63  
64 **Background:** Fluid overload, while common in the ICU and associated with serious sequelae, is hard to  
65 predict and may be influenced by ICU medication use. Machine learning (ML) approaches may offer  
66 advantages over traditional regression techniques to predict it. We compared the ability of traditional  
67 regression techniques and different ML-based modeling approaches to identify clinically meaningful fluid  
68 overload predictors.

69 **Methods:** This was a retrospective, observational cohort study of adult patients admitted to an ICU  $\geq$  72  
70 hours between 10/1/2015 and 10/31/2020 with available fluid balance data. Models to predict fluid  
71 overload (a positive fluid balance  $\geq$ 10% of the admission body weight) in the 48-72 hours after ICU  
72 admission were created. Potential patient and medication fluid overload predictor variables (n=28) were  
73 collected at either baseline or 24 hours after ICU admission. The optimal traditional logistic regression  
74 model was created using backward selection. Supervised, classification-based ML models were trained  
75 and optimized, including a meta-modeling approach. Area under the receiver operating characteristic  
76 (AUROC), positive predictive value (PPV), and negative predictive value (NPV) were compared between  
77 the traditional and ML fluid prediction models.

78 **Results:** A total of 49 of the 391 (12.5%) patients developed fluid overload. Among the ML models, the  
79 XGBoost model had the highest performance (AUROC 0.78, PPV 0.27, NPV 0.94) for fluid overload  
80 prediction. The XGBoost model performed similarly to the final traditional logistic regression model  
81 (AUROC 0.70; PPV 0.20, NPV 0.94). Feature importance analysis revealed severity of illness scores and  
82 medication-related data were the most important predictors of fluid overload.

83 **Conclusion:** In the context of our study, ML and traditional models appear to perform similarly to predict  
84 fluid overload in the ICU. Baseline severity of illness and ICU medication regimen complexity are  
85 important predictors of fluid overload.

86 **KEYWORDS:** critical care; fluid overload; prediction; medication regimen complexity; machine  
87 learning

88

## 89 INTRODUCTION

90 Fluid overload, a frequent and unintended consequence of the resuscitation process in critically ill adults  
91 may result in increased rates of acute kidney injury and invasive mechanical ventilation initiation,  
92 prolonged intensive care unit (ICU) stay, and mortality (1, 2). Timely de-resuscitation to remove excess  
93 fluid is associated with improved outcomes (3-6). While the predictors of volume responsiveness are  
94 well-established (7, 8), the predictors for ICU fluid overload remain unclear (7, 8). Development of  
95 rigorous fluid overload prediction algorithms could shorten the time to the implementation of fluid  
96 overload mitigation strategies [e.g., concentration of intravenous (IV) fluid products, discontinuation of  
97 maintenance fluids, administration of diuretics] and improve outcomes.

98 Non-diuretic ICU medication use may affect fluid overload risk; preliminary data suggests the  
99 medication regimen complexity-ICU (MRC-ICU) score is associated with both fluid overload and fluid  
100 balance (9). This score has also been shown to predict mortality and length of stay and also the  
101 medication interventions needed to optimize a patient's pharmacotherapy regimen (10-17). Therefore,  
102 quantifying patient-specific, medication-related data is likely an important consideration in the prediction  
103 of fluid overload in critically adults (2, 18, 19).

104 Event prediction in the ICU remains a perennial area of research given the many challenges that  
105 exist for clinicians to accurately predict clinical outcomes in the highly complex and dynamic critical care  
106 environment (20, 21). Artificial intelligence and machine learning techniques have been proposed as a  
107 method to improve ICU clinical outcome prediction given their unique ability to handle multi-  
108 dimensional problems and identify novel patterns within the vast troves of continuously-generated patient  
109 data (19, 22-24). However, to some ICU clinicians, the use of artificial intelligence/machine learning  
110 approaches to predict clinical events may have a 'black-box effect,' which can ultimately preclude  
111 implementation. The rigorous evaluation of whether artificial intelligence-based approaches predict  
112 clinical events better than traditional regression models (or clinical expertise alone) remains a key  
113 question in critical care practice (25-29).

114 In this study, we sought to compare the ability of machine learning approaches to traditional  
115 regression models to predict fluid overload and the individual predictors for its occurrence in critically ill  
116 adults. We hypothesized that advanced machine learning techniques perform better than traditional  
117 regression models to predict fluid overload and that the predictors for fluid overload identified through  
118 machine learning approaches may be different.

## 119 **METHODS**

120 We conducted a retrospective, observational study of adults admitted ICUs at the University of North  
121 Carolina Health System (UNCHS), an integrated health system, who had fluid overload data available.  
122 The protocol for this study was approved with waivers of informed consent and HIPAA authorization  
123 granted by UNHCS Institutional Review Board (approval number: (Project00001541); approval date:  
124 October 2021). Procedures followed in the study were in accordance with the ethical standards of the of  
125 the UNHCS Institutional Review Board and the Helsinki Declaration of 1975, as most recently amended  
126 (30). The reporting of this study adheres to the STrengthening and reporting of OBservational data in  
127 Epidemiology statement (31).

### 128 **Population**

129 A random sample of 1,000 adults ( $\geq 18$  years) admitted to an ICU at UNCHS between October 2015 and  
130 October 2020 was generated. Patients on their index ICU admission with fluid balance data available for  
131 the first 72 hours were included (**Supplemental Digital Content (SDC) Figure 1**). Patients were  
132 excluded if the admission was not their index ICU admission.

### 133 **Data Collection and Outcomes**

134 De-identified UNCHS electronic health record (EHR) data (Epic Systems, Verona, WI) housed in the  
135 Carolina Data Warehouse (CDW) was extracted by a trained CDW data analyst. The primary outcome  
136 was the presence of fluid overload at the 48-72 hours (i.e., day 3) after ICU admission. Fluid overload  
137 was defined as a positive fluid balance in milliliters (mL) greater than or equal to 10% of the patient's  
138 admission body weight in kilograms (kg) (2, 32). For example, a patient with a body weight of 100kg at  
139 ICU admission having a positive fluid balance at 72 hours of 12,000 mL (or 12kg) would be considered to

140 have fluid overload. A secondary outcome was the amount of fluid overload as a function of body weight.

141 For example, the aforementioned patient would have a fluid overload amount of 12%.

142 Following a literature review, and through investigator consensus, potential predictor variables

143 for fluid overload were defined (2, 33-36). A total of potential 28 predictors were identified: 1) *ICU*

144 *baseline*: age  $\geq 65$  years, sex, admission to a medical (vs. surgical) ICU, primary ICU admission

145 diagnosis (i.e., cardiac, chronic kidney disease, heart failure, hepatic, pulmonary, sepsis, trauma), and

146 select co-morbidities (i.e., chronic kidney disease, heart failure); 2) *24 hours after ICU admission*:

147 APACHE II and SOFA score (using worst values in the 24 hour period), use of supportive care devices

148 (i.e., renal replacement therapy, invasive mechanical ventilation), serum laboratory values (i.e., albumin

149  $< 3$  mg/dL, bicarbonate  $< 22$  mEq/L or  $> 29$  mEq/L, chloride  $\geq 110$  mEq/L, creatinine  $\geq 1.5$  mg/dL,

150 lactate  $\geq 2$  mmol/L, potassium  $\geq 5.5$  mEq/L, sodium  $\geq 148$  mEq/L or  $< 134$  mEq/L), fluid balance (mL),

151 and presence of acute kidney injury (as defined by need for renal replacement therapy or serum creatinine

152 greater than or equal two times baseline); 3) *Medication data at 24 hours*: MRC-ICU score, vasopressor

153 use in the first 24 hours, use of continuous medication infusions, and the number of continuous

154 medication infusions.

## 155 **Data Analysis**

### 156 *Data Missingness*

157 Due to the hypothesis-generating nature of our study and the lack of published data on ICU fluid overload

158 prediction, no attempt was made to estimate a study sample size. Multiple imputation (10) imputations

159 per variable was applied for all missing data (see **Supplemental Digital Content (SDC)**).

### 160 *Machine Learning Models*

161 We employed Random Forest, SVM and XGBoost for the task of modeling the presence of fluid overload

162 (37-39). During the model training on each of the ten imputed training sets, 5-fold cross validation was

163 applied for Random Forest, SVM and XGBoost to choose the hyperparameters for these machine learning

164 models. With the optimal hyperparameters, the models were fitted again on the corresponding imputed

165 training set. Predictions for probability of fluid overload were made on each of the ten imputed testing

166 sets using the corresponding optimal model. For Random Forest, two hyperparameters were tuned  
167 (number of trees and number of variables randomly sampled as candidates at each split). For SVM, linear  
168 kernel and cost of constraints violation were tuned. For XGBoost, two hyperparameters were tuned  
169 (maximum depth of a tree and maximum number of boosting iterations). For each model, there were ten  
170 different imputed test sets that then generated ten different predictions. These predictions of the  
171 probability for fluid overload were averaged as the final prediction.

172 For the degree of fluid overload, we built models with the amount of fluid overload at 72 hours.  
173 Since this is a continuous variable, we employed their regression of the above machine learning models:  
174 Random Forest regression, SVM regression, and XGBoost regression. For XGBoost, feature importance  
175 was measured as the frequency a feature was used in the trees. For Random Forest, feature importance  
176 was measured by mean decrease in node impurity. Because ten different models were used on each  
177 imputed dataset, ten different feature importance lists were generated for each. A subsequent analysis  
178 modeling fluid overload as a continuous variable (percent of net milliliters of fluid by body weight)  
179 instead of dichotomous presence or absence of fluid overload) was performed (see **SDC**).

#### 180 *Traditional Regression Models*

181 991After multiple imputation, each of the ten completed datasets was split into training data and testing  
182 data using an 80:20 ratio. Subsequently, a full logistic regression model was built for the presence of fluid  
183 overload for each of the ten complete training sets. We then applied backward elimination to select the  
184 final model. The initial set of variables for the variable selection were determined by the significance of  
185 variables in the ten full models. We built our linear regression models so that the degree of fluid overload  
186 was similar to that of the ten completed training sets. In order to compare these models with the MRC-  
187 ICU only model, we also built logistic regression and linear regression models with MRC-ICU as the sole  
188 predictor in the ten training sets. After model fitting, model fits were pooled using Rubin's method (40).  
189 Using the pooled models, odds ratios (OR) and their 95% confidence intervals (CI) were reported.

## 190 **RESULTS**

191 A total of 49 (12.5%) of the 391 included patients had fluid overload on ICU day 3. The degree of  
192 day 3 fluid overload was significantly greater in the fluid overload (vs non overload) patients (16.6% vs  
193 2.2%,  $p < 0.01$ ). Overall, the mean APACHE II score was  $15.7 \pm 6.6$ , mean SOFA score was  $8.3 \pm 3.3$ ,  
194 and MRC-ICU score was  $11.8 \pm 8.7$ . A significantly greater proportion of fluid overload patients (vs.  
195 those without) had an elevated serum lactate  $\geq 2$  mmol/L (32.7% vs. 14.9%,  $p = 0.01$ ) and AKI (28.6%  
196 vs. 10.5%,  $p < 0.001$ ) at 24 hours and positive fluid balance (1,840 mL vs. 390 mL,  $p < 0.001$ ) on ICU  
197 day 3. All model covariates are summarized in **Table 1**. At ICU day 3, patients with fluid overload (vs  
198 those without) were more likely to be dead (20.4% vs. 7.3%,  $p = 0.01$ ), have AKI (34.7% vs. 15.8%,  $p <$   
199  $0.001$ ), and remain on mechanical ventilation (12.7% vs. 4.2%,  $p = 0.05$ ).

200 Among the machine learning models, XGBoost demonstrated the highest AUROC (0.78)  
201 compared to SVM (0.69) and RF (0.76) and was associated with a PPV of 0.27 and NPV of 0.94.  
202 Notably, all models tested at relatively poor PPV. In comparison, stepwise logistic regression had an  
203 AUROC of 0.70, PPV 0.26, and NPV 0.94. Full results are reported in **Table 2**, and AUROC curves for  
204 all models are provided in **SDC Supplemental Figure 2**. Results of the full logistic regression are  
205 reported in **SDC Supplemental Table 1**. Stepwise regression resulted in a more parsimonious model (7  
206 variables vs. 31 variables) but demonstrated similar performance to the machine learning models (**SDC**  
207 **Supplementary Table 2**). In the stepwise regression, presence of sepsis, male sex, the SOFA score at 24  
208 hours, and the 24 hour serum sodium and bicarbonate comprised the stepwise regression model (**Table 2**).  
209 In an analysis of MRC-ICU as a single predictor for fluid overload, the model had an AUROC of 0.74  
210 (0.60-0.84), sensitivity 0.62 (0.35-0.85), specificity 0.70 (0.63-0.77), PPV 0.16 (0.08-0.27), and NPV  
211 0.96 (0.90-0.98).

212 Feature importance graphs were plotted for XGBoost (**Figure 1**), RF (**SDC Supplemental**  
213 **Figure 3**) and SVM (**SDC 5 Supplemental Figure 4**). Among the 10 different feature importance lists  
214 generated for each model, differences between top features were noted. For example, for two of the  
215 machine learning models, XGBoost (**Figure 2**) and RF, the top five most important features were fluid  
216 balance at 24 hours, SOFA score at 24 hours, MRC-ICU at 24 hours, APACHE II at 24 hours, and the



217 number of continuous infusions at 24 hours. While the stepwise regression model found fluid balance at  
218 24 hours and APACHE II at 24 hours to be top features, the SOFA score at 24 hours, the MRC-ICU at 24  
219 hours and the number of continuous infusions were not found to be model features.

220 The full regression results for predicting the amount of fluid overload at 72 hours are reported in  
221 **SDC Supplemental Table 3**. For stepwise regression, twelve variables were included with fluid balance,  
222 laboratory values, and severity of illness being significant predictors (**SDC Supplemental Table 4**). All  
223 models demonstrated similar performance as measured by MSE (**SDC Supplemental Table 5**). Feature  
224 importance graphs are presented in **SDC Supplemental Figures 5-7**).

## 225 **DISCUSSION**

226 Although machine learning models have been shown to outperform traditional regression models in a  
227 variety of settings (41, 42), the potential benefits of machine learning in critical care remain an open field  
228 of exploration, in part due to a current lack of rigorous comparison in high quality ICU datasets (27, 43,  
229 44). Our analysis represents the first published comparison of machine learning approaches with  
230 traditional regression methods to predict fluid overload using a novel dataset with granular medication  
231 data.

232 We report that machine learning and logistic regression analyses demonstrate a similar predictive  
233 power to identify patients with fluid overload on day 3 of their ICU stay. Although use of machine  
234 learning did not appear to improve predictive performance over regression analysis, it expanded the  
235 number of variables critical to fluid overload prediction and highlights the importance of further artificial  
236 intelligence-based exploration in this area. This analysis of individual predictors may help bedside  
237 clinicians better understand how the machine learning models work and may help overcome their ‘black  
238 box’ hesitancy to trust machine learning-generated results (45, 46). For example, feature importance  
239 graphs for the machine learning analyses found complexity of the daily ICU medication regimen (i.e.,  
240 MRC-ICU score), which includes the number of intravenous medication infusions (the primary method to  
241 administer medications in this population and a primary source of fluids to a patient), to be an important

242 contributor to fluid overload. In comparison, in the traditional multivariable regression, the MRC-ICU  
243 score was not associated with fluid overload. This may be because machine learning analyses better  
244 account for severity of illness and the response of clinicians to respond to this severity by administering  
245 more medication infusions leading to a more complex daily medication regimen; however, the methods  
246 applied, including feature importance, preclude causal inference at this juncture. As such, our results  
247 highlight the unique power of machine learning to identify complex relationships that can be further  
248 elucidated via machine-learning based causal inference modeling and other designs aimed at causation (2,  
249 18).

250         Optimizing fluid management (or fluid stewardship) has been previously defined by the ROSE  
251 model of Resuscitation, Optimization, Stabilization, and dE-resuscitation (33). After an initial 24-48 hour  
252 period characterized by overt volume resuscitation (e.g., a crystalloid bolus) and IV medication initiation  
253 (e.g., antibiotics), and the associated fluid administration, the care priority shifts from volume  
254 administration to volume removal. While comprehensive fluid stewardship management strategies  
255 including reduced fluid use and diuretic administration can effectively reduce fluid overload and its  
256 sequelae, they are often deployed too late (1, 2). Interestingly, some reports have indicated ‘hidden fluids’  
257 (defined as blood products, enteral nutrition, flushes, and intravenous medications) were significantly  
258 associated with the development of fluid overload. While in critical illness many of these ‘hidden fluids’  
259 are necessary (e.g., blood products), given that intravenous medications account for over 40% of total  
260 fluid intake in this analysis, interventions such as concentrating intravenous medications, employing oral  
261 formulations when feasible, careful evaluation of maintenance fluids, and antibiotic de-escalation are  
262 potentially still viable even in high illness severity that can reduce this complication. However, weighing  
263 risks and benefits associated with these interventions in context may yet be aided by more quantitative  
264 prediction data (50, 51). Overall, de-resuscitation and fluid stewardship can be deceptively complex (47).  
265 In a patient with shock, balancing the dueling forces of volume responsiveness assessment and timely  
266 volume resuscitation with the risks associated with fluid overload represents a highly complex Goldilocks

267 scenario that requires clinicians to have high clinical precision, essentially pivoting ‘on a dime’, from a  
268 strategy of aggressive volume expansion to one of rapid volume removal (34, 48, 49).

269         Despite the complexities of this decision process, limited prediction tools for fluid overload are  
270 available to assist clinicians at the ICU bedside. As such, real-time recognition identifying when to make  
271 the shift from resuscitation to de-resuscitation has the potential to improve bedside management.  
272 However, to go beyond the hourly assessment of ‘Ins and Outs’ would require accurate prediction of  
273 future fluid overload risk and the adverse events associated with it, in the time-dependent context of  
274 intervention delivery (e.g., diuretics). In such a scenario, an algorithm would be able to accurately  
275 interpret a septic patient who is 3 liters positive 24 hours after fluid resuscitation initiation as being in a  
276 ‘green zone’ (i.e., appropriately resuscitated). However, 24 hours later, if the same patient is 4 liters  
277 positive while off vasopressors and with down-trending sepsis markers the algorithm could alert  
278 clinicians that the patient is now in a ‘yellow zone’ where interventions like diuretic therapy and fluid  
279 reductions are required to reduce acute kidney injury and intubation risk. This type of real-time predictive  
280 capability could support continuous clinician decision-making but requires evaluation outside the scope of  
281 our current study.

282         Fluid overload also presents an important test case for exploring and adapting artificial  
283 intelligence methods to ICU problems, particularly those related to ICU medication use. Fluid overload  
284 represents a uniquely *intervenable event* in the ICU. Intervenable events share three key characteristics:  
285 they are *predictable*, *preventable*, and otherwise associated with *poor* outcomes. The results of our study,  
286 and others, indicate that fluid overload can be *predicted* with modeling of some kind, especially given its  
287 ability to be quantitatively defined (50-52). Fluid overload has been associated with poor outcomes  
288 including acute kidney injury, delirium, poor respiratory outcomes, prolonged length of stay, and  
289 potentially increasing mortality (2, 35, 53-56). Evidence demonstrates the timely recognition and  
290 management of fluid overload is feasible and is associated with reduced mortality and time in the ICU (3,  
291 57-60). Notably, fluid stewardship has been adapted by critical care pharmacists as key component of

292 comprehensive medication management (5, 6, 60). As such, these results may support other  
293 investigations as they identify patients in whom it is safe to initiate de-resuscitation or importantly never  
294 needed that degree of fluid volume initially and at the bedside may prompt clinicians to be more targeted  
295 in therapies initiated or aggressive in curtailing early ‘hidden’ fluids to avoid the complications of fluid  
296 overload and/or the need for a highly interventional period of de-resuscitation (e.g., diuretics, dialysis).  
297 Artificial intelligence may be particularly well suited to bolster these efforts, and thus while feature  
298 importance analyses cannot provide foundation for causal inference, they may guide such future  
299 investigations.

300 Our study has limitations. Our patient sample may have been too small to demonstrate superiority  
301 of the machine learning approaches compared to traditional regression, and no validation in a separate,  
302 external dataset was undertaken at this juncture (61). Bias may exist due to which patients had fluid  
303 balance data available. Other predictors for fluid overload not included in our models may exist (62). By  
304 relying on prediction data derived in the first 24 hours of ICU admission, we did not fully capture the  
305 dynamic nature of critical illness over the entire three day ICU period before fluid overload occurred.  
306 Future time-dependent evaluations of changing features employing unsupervised learning techniques may  
307 yield novel insights.

## 308 **CONCLUSION**

309 Fluid overload is an important, intervenable event in the ICU population. Incorporation of medication-  
310 related variables and artificial intelligence has demonstrated promise to improve prediction that may  
311 ultimately guide timely intervention and mitigation of this ICU complication; however, comparative  
312 advantages over traditional modeling techniques may remain warranted.

313

314

315 **Declarations**

316

317 **Ethical Approval**

318

319 The protocol for this study was approved with waivers of informed consent and HIPAA authorization  
320 granted by UNHCS Institutional Review Board (approval number: (Project00001541); approval date:  
321 October 2021).

322

323 **Author Contributions**

324

325 A.S. was responsible for project execution, design, and initial manuscript writing. J.D., D.M.,  
326 and R.K. provided critical revisions of manuscript, data interpretation, and senior level oversight.  
327 M.Y., T.Z, and X.C. handled data pre-processing and analysis (M.Y., T.Z.) and methodology  
328 support and data interpretation (X.C., R.K.). B.M. served as site coordinator for all data  
329 validation and procurement as well as manuscript revisions and data interpretation. S.S., M.B.,  
330 and S.R. provided clinical interpretation, results interpretation, and manuscript revision.

331

332 **Availability of data & materials**

333

334 The datasets used and/or analyzed during the current study available from the corresponding  
335 author on reasonable request.

336

337

338

339

340

341

342 **References**

- 343
- 344 1. Carr JR, Hawkins WA, Newsome AS, et al: Fluid stewardship of maintenance Intravenous  
345 fluids. *J Pharm Pract* 2021:8971900211008261
- 346
- 347 2. Hawkins WA, Smith SE, Newsome AS, et al: Fluid stewardship during critical illness: a  
348 call to action. *J Pharm Pract* 2020; 33(6):863-873
- 349
- 350 3. Bissell BD, Laine ME, Thompson Bastin ML, et al: Impact of protocolized diuresis for de-  
351 resuscitation in the intensive care unit. *Crit Care* 2020; 24(1):70
- 352
- 353 4. Jones TW, Chase AM, Bruning R, et al: Early diuretics for de-resuscitation in septic  
354 patients with left ventricular dysfunction. *Clin Med Insights Cardiol* 2022;  
355 16:11795468221095875
- 356
- 357 5. Hawkins WA, Butler SA, Poirier N, et al: From theory to bedside: implementation of fluid  
358 stewardship in a medical ICU pharmacy practice. *Am J Health Syst Pharm* 2022 Jun  
359 7;79(12):984-992.
- 360
- 361 6. Bissell BD, Donaldson JC, Morris PE, et al: A narrative review of pharmacologic de-  
362 resuscitation in the critically ill. *J Crit Care* 2020; 59:156-162
- 363
- 364 7. Messmer AS, Moser M, Zuercher P, et al: Fluid overload phenotypes in critical illness-a  
365 machine learning approach. *J Clin Med* 2022; 11(2)
- 366
- 367 8. Zhang Z, Ho KM, Hong Y: Machine learning for the prediction of volume responsiveness  
368 in patients with oliguric acute kidney injury in critical care. *Crit Care* 2019; 23(1):112
- 369
- 370 9. Olney WJ, Chase AM, Hannah SA, et al: Medication regimen complexity score as an  
371 indicator of fluid balance in critically ill Patients. *J Pharm Pract* 2021:897190021999792
- 372
- 373 10. Sikora A, Ayyala D, Rech MA, et al: Impact of pharmacists to improve patient care in the  
374 critically ill: a large multicenter analysis using meaningful metrics with the medication  
375 regimen complexity-ICU (MRC-ICU) score. *Crit Care Med* 2022; 50(9):1318-1328
- 376
- 377 11. Newsome AS, Smith SE, Olney WJ, et al: Multicenter validation of a novel medication-  
378 regimen complexity scoring tool. *Am J Health Syst Pharm* 2020; 77(6):474-478
- 379
- 380 12. Newsome AS, Anderson D, Gwynn ME, et al: Characterization of changes in medication  
381 complexity using a modified scoring tool. *Am J Health Syst Pharm* 2019;  
382 76(Supplement\_4):S92-s95
- 383
- 384 13. Gwynn ME, Poisson MO, Waller JL, et al: Development and validation of a medication  
385 regimen complexity scoring tool for critically ill patients. *Am J Health Syst Pharm* 2019;  
386 76(Supplement\_2):S34-s40

- 387 Al-Mamun MA, Brothers T, Newsome AS: Development of Machine Learning Models to  
388 Validate a Medication Regimen Complexity Scoring Tool for Critically Ill Patients. *Ann*  
389 *Pharmacother* 2021; 55(4):421-429  
390
- 391 Smith SE, Shelley R, Sikora A: Medication regimen complexity vs patient acuity for  
392 predicting critical care pharmacist interventions. *Am J Health Syst Pharm* 2022;  
393 79(8):651-655  
394
- 395 Webb AJ, Rowe S, Newsome AS: A descriptive report of the rapid implementation of  
396 automated MRC-ICU calculations in the EMR of an academic medical center. *Am J*  
397 *Health Syst Pharm* 2022; 79(12):979-983  
398
- 399 Newsome AS, Smith SE, Olney WJ, et al: Medication regimen complexity is associated  
400 with pharmacist interventions and drug-drug interactions: A use of the novel MRC-ICU  
401 scoring tool. *Journal of the American College of Clinical Pharmacy*. 2020; 3(1):47-56  
402
- 403 Sanchez P, Voisey JP, Xia T, et al: Causal machine learning for healthcare and precision  
404 medicine. *R Soc Open Sci* 2022; 9(8):220638  
405
- 406 Iwase S, Nakada TA, Shimada T, et al: Prediction algorithm for ICU mortality and length  
407 of stay using machine learning. *Sci Rep* 2022; 12(1):12912  
408
- 409 Beil M, Sviri S, Flaatten H, et al: On predictions in critical care: The individual  
410 prognostication fallacy in elderly patients. *J Crit Care* 2021; 61:34-38  
411
- 412 Lovejoy CA, Buch V, Maruthappu M: Artificial intelligence in the intensive care unit.  
413 *Crit Care* 2019; 23(1):7  
414
- 415 Gutierrez G: Artificial Intelligence in the Intensive Care Unit. *Crit Care* 2020; 24(1):101  
416
- 417 Goh KH, Wang L, Yeow AYK, et al: Artificial intelligence in sepsis early prediction and  
418 diagnosis using unstructured data in healthcare. *Nat Commun* 2021; 12(1):711  
419
- 420 Hyland SL, Faltys M, Huser M, et al: Early prediction of circulatory failure in the  
421 intensive care unit using machine learning. *Nat Med* 2020; 26(3):364-373  
422
- 423 DeGrave AJ, Janizek JD, Lee SI: AI for radiographic COVID-19 detection selects  
424 shortcuts over signal. *medRxiv* 2020  
425
- 426 Nguyen D, Ngo B, vanSonnenberg E: AI in the Intensive Care Unit: Up-to-Date Review. *J*  
427 *Intensive Care Med* 2021; 36(10):1115-1123  
428
- 429 Yoon JH, Pinsky MR, Clermont G: Artificial Intelligence in Critical Care Medicine. *Crit*  
430 *Care* 2022; 26(1):75



- 431 28. Farion KJ, Wilk S, Michalowski W, et al: Comparing predictions made by a prediction  
432 model, clinical score, and physicians: pediatric asthma exacerbations in the emergency  
433 department. *Appl Clin Inform* 2013; 4(3):376-391  
434
- 435 29. Feng JZ, Wang Y, Peng J, et al: Comparison between logistic regression and machine  
436 learning algorithms on survival prediction of traumatic brain injuries. *J Crit Care* 2019;  
437 54:110-116  
438
- 439 30. World Medical Association: World Medical Association Declaration of Helsinki: Ethical  
440 principles for medical research involving human subjects. *JAMA* 2023; 312:2191-2194.  
441
- 442 31. Von Elm E AD, Egger M, et al: STROBE Initiative: Strengthening the reporting of  
443 observational studies in epidemiology (STROBE) statement: Guidelines for reporting  
444 observational studies. *BMJ* 2007; 335:806-808.  
445
- 446 32. Bouchard J, Soroko SB, Chertow GM, et al: Fluid accumulation, survival and recovery of  
447 kidney function in critically ill patients with acute kidney injury. *Kidney Int* 2009;  
448 76(4):422-427  
449
- 450 33. Carr JR, Hawkins WA, Newsome AS, et al: Fluid Stewardship of Maintenance  
451 Intravenous Fluids. *J Pharm Pract* 2022; 35(5):769-782  
452
- 453 34. Malbrain M, Van Regenmortel N, Saugel B, et al: Principles of fluid management and  
454 stewardship in septic shock: it is time to consider the four D's and the four phases of fluid  
455 therapy. *Ann Intensive Care* 2018; 8(1):66  
456
- 457 35. Claire-Del Granado R, Mehta RL: Fluid overload in the ICU: evaluation and  
458 management. *BMC Nephrol* 2016; 17(1):109  
459
- 460 36. O'Connor ME, Prowle JR: Fluid Overload. *Crit Care Clin* 2015; 31(4):803-821  
461
- 462 37. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: Proceedings of the  
463 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining  
464 [Internet]. New York, NY, USA: ACM; 2016. p. 785–94. Available from:  
465 <http://doi.acm.org/10.1145/2939672.2939785>  
466
- 467 38. Cortes C, Vapnik V. Support-vector networks. *Machine learning*. 1995;20(3):273–97.  
468
- 469 39. Ho TK. Random decision forests. In: Proceedings of 3rd international conference on  
470 document analysis and recognition. 1995. p. 278–82.  
471
- 472 40. Rubin DB: Multiple imputation for nonresponse in surveys. Hoboken, N.J. ;, Wiley-  
473 Interscience, 2004  
474
- 475 41. Topol EJ: Deep medicine : how artificial intelligence can make healthcare human again.  
476 First edition. ed. New York, Basic Books,, 2019, p^pp 1 online resource



- 427 Kahneman D, Sibony O, Sunstein CR: Noise : a flaw in human judgment. First edition.  
478 Edition. New York, Little, Brown Spark, 2021  
479
- 480 D'Hondt E, Ashby TJ, Chakroun I, et al: Identifying and evaluating barriers for the  
481 implementation of machine learning in the intensive care unit. *Commun Med (Lond)* 2022;  
482 2(1):162  
483
- 484 van de Sande D, van Genderen ME, Huiskens J, et al: Moving from bytes to bedside: a  
485 systematic review on the use of artificial intelligence in the intensive care unit. *Intensive*  
486 *Care Med* 2021; 47(7):750-760  
487
- 488 Moss L, Corsar D, Shaw M, et al: Demystifying the Black Box: The Importance of  
489 Interpretability of Predictive Models in Neurocritical Care. *Neurocrit Care* 2022;  
490 37(Suppl 2):185-191  
491
- 492 The Lancet Respiratory M: Opening the black box of machine learning. *Lancet Respir*  
493 *Med* 2018; 6(11):801  
494
- 495 Malbrain M, Martin G, Ostermann M: Everything you need to know about deresuscitation.  
496 *Intensive Care Med* 2022; 48(12):1781-1786  
497
- 498 Gelbart B, Serpa Neto A, Stephens D, et al: Fluid Accumulation in Mechanically  
499 Ventilated, Critically Ill Children: Retrospective Cohort Study of Prevalence and  
500 Outcome. *Pediatr Crit Care Med* 2022; 23(12):990-998  
501
- 502 National Heart L, Blood Institute Acute Respiratory Distress Syndrome Clinical Trials N,  
503 Wiedemann HP, et al: Comparison of two fluid-management strategies in acute lung  
504 injury. *N Engl J Med* 2006; 354(24):2564-2575  
505
- 506 Gamble KC, Smith SE, Bland CM, et al: Hidden Fluids in Plain Sight: Identifying  
507 Intravenous Medication Classes as Contributors to Intensive Care Unit Fluid Intake. *Hosp*  
508 *Pharm* 2022; 57(2):230-236  
509
- 510 Branam T, Smith SE, Newsome AS, et al: Association of hidden fluid administration with  
511 development of fluid overload reveals opportunities for targeted fluid minimization. *SAGE*  
512 *Open Med* 2020; 8:2050312120979464  
513
- 514 Mitchell KH, Carlbom D, Caldwell E, et al: Volume Overload: Prevalence, Risk Factors,  
515 and Functional Outcome in Survivors of Septic Shock. *Ann Am Thorac Soc* 2015;  
516 12(12):1837-1844  
517
- 518 Ouchi A, Sakuramoto H, Hoshino H, et al: Association between fluid overload and  
519 delirium/coma in mechanically ventilated patients. *Acute Med Surg* 2020; 7(1):e508  
520
- 521 Murphy CV, Schramm GE, Doherty JA, et al: The importance of fluid management in  
522 acute lung injury secondary to septic shock. *Chest* 2009; 136(1):102-109

- 523 55. Boyd JH, Forbes J, Nakada TA, et al: Fluid resuscitation in septic shock: a positive fluid  
524 balance and elevated central venous pressure are associated with increased mortality. *Crit*  
525 *Care Med* 2011; 39(2):259-265  
526
- 527 56. Woodward CW, Lambert J, Ortiz-Soriano V, et al: Fluid Overload Associates With Major  
528 Adverse Kidney Events in Critically Ill Patients With Acute Kidney Injury Requiring  
529 Continuous Renal Replacement Therapy. *Crit Care Med* 2019; 47(9):e753-e760  
530
- 531 57. Hawkins WA, Butler SA, Poirier N, et al: From theory to bedside: Implementation of fluid  
532 stewardship in a medical ICU pharmacy practice. *Am J Health Syst Pharm* 2022;  
533 79(12):984-992  
534
- 535 58. Silversides JA, Perner A, Malbrain M: Liberal versus restrictive fluid therapy in critically  
536 ill patients. *Intensive Care Med* 2019; 45(10):1440-1442  
537
- 538 59. Goldstein S, Bagshaw S, Cecconi M, et al: Pharmacological management of fluid  
539 overload. *Br J Anaesth* 2014; 113(5):756-763  
540
- 541 60. Silversides JA, McAuley DF, Blackwood B, et al: Fluid management and deresuscitation  
542 practices: A survey of critical care physicians. *J Intensive Care Soc* 2020; 21(2):111-118  
543
- 544 61. Burkov A: The hundred-page machine learning book. Quebec City, Canada, Andriy  
545 Burkov, 2019  
546
- 547 62. Qin X, Zhang W, Hu X, et al: A deep learning model to identify the fluid overload status  
548 in critically ill patients based on chest X-ray images. *Pol Arch Intern Med* 2023; 133(2)  
549

550 **Figure 1.** Feature importance for presence of fluid overload prediction with XGBoost

551

552 **Figure 2.** Most common features for presence of fluid overload prediction with XGBoost imputations

553

554

555

556

557 **Table 1.** Study cohort characteristics by presence of fluid overload within 72 hours of ICU admission

	All (n = 391)	Fluid Overload (n = 49)	No Fluid Overload (n = 342)	p-value
<b>ICU Baseline</b>				
Age ≥ 65 years	202 (51.7)	19 (38.8)	183 (53.5)	0.08
Male sex	213 (54.5)	23 (46.9)	190 (55.6)	0.33
<b>Chronic comorbidities</b>				
Chronic kidney disease	13 (3.3)	1 (2.0)	12 (3.5)	0.06
Heart failure	19 (4.9)	2 (4.1)	17 (4.9)	0.06
Admission to medical ICU	156 (39.9)	24 (48.9)	132 (38.6)	0.22
<b>Primary ICU Admission Diagnosis</b>				
Cardiac	81 (20.7)	3 (6.1)	78 (22.8)	0.06
Chronic kidney disease	13 (3.3)	1 (2.0)	12 (3.5)	
Hepatic	6 (1.5)	1 (2.0)	5 (1.5)	
Pulmonary	58 (14.8)	8 (16.3)	50 (14.6)	
Sepsis/septic shock	29 (7.4)	7 (14.3)	22 (6.4)	
Trauma	10 (2.6)	3 (6.1)	7 (2.0)	
<b>24 hours after ICU admission</b>				
<b>Severity of illness, mean (SD)</b>				
APACHE II Score	15.7 (6.6)	17.5 (7.0)	15.4 (6.6)	0.06
SOFA Score	8.3 (3.3)	9.9 (4.6)	8.2 (3.1)	0.07
<b>Supportive devices</b>				
Any renal replacement therapy	5 (1.3)	1 (2.0)	4 (1.2)	1.00
Any mechanical ventilation	140 (35.8)	21 (42.9)	119 (34.8)	0.53
<b>Serum laboratory values</b>				
Albumin <3 mg/dL	88 (22.5)	18 (36.7)	70 (20.5)	0.02
Bicarbonate < 22 mEq/L	74 (18.9)	14 (28.6)	60 (17.5)	0.16
Bicarbonate > 29 mEq/L	64 (16.4)	6 (12.2)	58 (16.9)	
Creatinine ≥ 1.5 mg/dL	28 (7.2)	7 (14.3)	21 (6.1)	0.02
Chloride ≥ 110 mEq/L	125 (31.9)	19 (38.8)	106 (30.9)	0.33
Potassium ≥ 5.5 mEq/L	19 (4.9)	5 (10.2)	14 (4.1)	0.12
Lactate ≥ 2 mmol/L	67 (17.1)	16 (32.7)	51 (14.9)	0.01
Sodium ≥ 148 mEq/L	22 (5.6)	6 (12.2)	16 (4.7)	0.01
Sodium <134 mEq/L	33 (8.4)	4 (8.1)	29 (8.5)	
Fluid balance (mL), mean (SD)	570 (1960)	1840 (301)	390 (168)	<0.001
Acute kidney injury	50 (12.8)	14 (28.6)	26 (10.5)	< 0.001
<b>Medications</b>				
MRC-ICU, mean (SD)	11.8 (8.7)	13.4 (8.4)	11.5 (8.7)	0.06
Any vasopressor	119 (30.4)	16 (32.6)	103 (30.1)	0.85
Any continuous infusions	249 (63.6)	34 (69.3)	215 (62.8)	0.47
Infusions / patient, mean (SD)	2.29 (3.3)	1.98 (2.2)	2.33 (3.4)	0.35

558 Data are presented as n (%) unless otherwise stated

559

560

561

562

563

564

565

566

567 **Table 2.** Performance of presence of fluid overload prediction models, mean (confidence interval)  
 568

	<b>AUROC</b>	<b>Sensitivity</b>	<b>Specificity</b>	<b>PPV</b>	<b>NPV</b>
<i>Traditional regression</i>					
<b>All variables</b>	0.70 (0.53, 0.82)	0.43 (0.19, 0.70)	0.85 (0.79, 0.89)	0.20 (0.08, 0.37)	0.94 (0.89, 0.97)
<b>Stepwise Selected Regression</b>	0.70 (0.52, 0.82)	0.43 (0.19, 0.70)	0.89 (0.84, 0.93)	0.26 (0.11, 0.47)	0.94 (0.90, 0.97)
<i>Supervised machine learning models</i>					
<b>Random Forest</b>	0.76 (0.62, 0.86)	0.56 (0.29, 0.80)	0.8571 (0.80, 0.90)	0.25 (0.12, 0.43)	0.95 (0.91, 0.98)
<b>Support Vector Machine</b>	0.69 (0.51, 0.82)	0.50 (0.24, 0.75)	0.82 (0.76, 0.88)	0.21 (0.09, 0.36)	0.94 (0.90, 0.97)
<b>XGBoost</b>	0.78 (0.62, 0.87)	0.37 (0.15, 0.64)	0.91 (0.86, 0.94)	0.27 (0.10, 0.50)	0.94 (0.89, 0.97)

569 AUROC: area under the receiver operating characteristic; PPV: positive predictive value; NPV: negative predictive  
 570 value  
 571  
 572  
 573  
 574  
 575  
 576  
 577