Original Study

ANNALS OF SURGERY OPEN

OPEN

Using Artificial Neural Networks to Predict Intra-Abdominal Abscess Risk Post-Appendectomy

Morouge M. Alramadhan, MD,* Hassan S. Al Khatib, MSc,* James R. Murphy, PhD,* KuoJen Tsao, MD,† and Michael L. Chang, MD*

Downloaded from http://journals.lww.com/aosopen by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 07/10/2023

Objective: To determine if artificial neural networks (ANN) could predict the risk of intra-abdominal abscess (IAA) development post-appendectomy.

Background: IAA formation occurs in 13.6% to 14.6% of appendicitis cases with "complicated" appendicitis as the most common cause of IAA. There remains inconsistency in describing the severity of appendicitis with variation in treatment with respect to perforated appendicitis.

Methods: Two "reproducible" ANN with different architectures were developed on demographic, clinical, and surgical information from a retrospective surgical dataset of 1574 patients less than 19 years old classified as either negative (n = 1,328) or positive (n = 246) for IAA post-appendectomy for appendicitis. Of 34 independent variables initially, 12 variables with the highest influence on the outcome selected for the final dataset for ANN model training and testing.

Results: A total of 1574 patients were used for training and test sets (80%/20% split). Model 1 achieved accuracy of 89.84%, sensitivity of 70%, and specificity of 93.61% on the test set. Model 2 achieved accuracy of 84.13%, sensitivity of 81.63%, and specificity of 84.6%.

Conclusions: ANN applied to selected variables can accurately predict patients who will have IAA post-appendectomy. Our reproducible and explainable ANNs potentially represent a state-of-the-art method for optimizing post-appendectomy care.

Keywords: artificial intelligence, intraabdominal abscess, pediatric

INTRODUCTION

Approximately 70,000 children are diagnosed with appendicitis each year in the United States, with up to 20% to 35% suffering from appendiceal perforation, especially in younger children. Intra-abdominal abscess (IAA) formation, the most common serious complication occurs in 13.6% to 14.6% of appendicitis cases.[1] Correspondingly, "complicated" appendicitis is the most common cause of IAA in children.[2] However, there remains inconsistency among surgeons in describing the severity of appendicitis,[3] though perforation is generally defined as a visible hole in the appendix or a free fecalith in the abdomen.[4] As controversy and variation in treatment (including nonoperative management) exist with respect to perforated appendicitis, predicting which appendicitis (perforated and nonperforated) patients who will develop complications would be beneficial.

In recent years, artificial neural networks (ANN) have become a tool in complicated clinical scenarios to help predict variable outcomes, including risk of acquiring infectious diseases.[5–8] ANN uses in health care have been established and discussed for various predictive tasks such as cancer prediction, surgery prognosis, clinical diagnosis, image analysis and interpretation, and drug development.[9–13] Compared with traditional machine learning models, ANN may be more sustainable over time.[13] An ANN is a system that is composed of interconnecting parallel nonlinear elements "artificial neurons" which are a type of mathematical function with limited numbers of inputs and outputs. Each artificial neuron processes input "signals," computes the signals by a nonlinear function, and produces an output.[14,15] The individual artificial neurons are connected via edges (the connection) to other individual artificial neurons. In this way, the ANN provide inputs and outputs to each other similar to biological neural networks from which ANN are inspired. The ANN can "learn" to perform tasks as a form of artificial intelligence by adjusting the weight of each connection and artificial neuron, highlighting certain neurons more than others within the ANN. Furthermore, the artificial neurons are typically arranged into layers, with the input layer receiving external data, passing signals to the next layer, with a final output layer that accomplishes the task assigned to the ANN.[14,15]

On reviewing the complexity of diagnosis of complicated and/or perforated appendicitis, coupled with variation in treatment approaches, our aim in this project is to develop a nonlinear predictive ANN model using various input variables to identify which patients are at risk for developing IAA following appendectomy for appendicitis. Application of this

From the *Division of Infectious Diseases, Department of Pediatrics, UTHealth Houston McGovern Medical School, Houston, TX; and †Division of General and Thoracic Pediatric Surgery, Department of Pediatric Surgery, UTHealth Houston McGovern Medical School, Houston, TX.

This manuscript was previously published as an abstract in a supplement to Open Forum Infect Dis: doi: 10.1093/ofid/ofz360.2116.

Disclosure: The authors declare that they have nothing to disclose.

All authors have contributed significantly to the work, have seen and approved the manuscript submission. M.M.A., H.S.A.K., M.L.C. had full access to the data in the study and take responsibility for the integrity of the data and accuracy of the results. M.M.A., H.S.A.K., J.M., K.T., M.L.C. contributed to the study concept and design. H.S.A.K. contributed to the development of artificial neural networks. M.M.A. contributed to the drafting of the manuscript.

Reprints: Morouge M. Alramadhan, MD, Division of Infectious Diseases, Department of Pediatrics, UTHealth Houston McGovern Medical School, 6431 Fannin St., MSB 3.126, Houston, TX 77030. Email: morouge.m.alramadhan@uth.tmc.edu.

Copyright © 2022 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Annals of Surgery Open (2022) 2:e168

Received: 27 September 2021; Accepted 18 April 2022

Published online 19 May 2022

DOI: 10.1097/AS9.0000000000000168

type of nonlinear predictive model to facilitate identification of the complication of IAA after diagnosis of appendicitis may help pediatric surgeons standardize postoperative care, optimize antibiotic use for appendicitis patients, and decrease hospital lengths of stay.

METHODS

Population

A total of 1594 patients were classified by our pediatric surgery faculty according to the American College of Surgeons National Surgical Quality Improvement Project Pediatric Data Definitions Committee[16] as negative (n = 1,328) or positive (n = 246) for IAA post-appendectomy for acute appendicitis within an internal institutional research database containing 6127 patients less than 19 years old who had an appendectomy at our single center between 2009 and 2018. Our center is an academic teaching facility that serves as a tertiary referral center for a large urban metropolitan area. All pediatric appendectomy procedures, whether open or laparoscopic, will have resident or fellow trainees and pediatric surgery faculty present 24 hours a day, 365 days a year. Pediatric appendectomies are exclusively staffed by faculty from the Division of General and Thoracic Pediatric Surgery who all are certified for pediatric surgery by the American Board of Surgery.

From the 1594 patients that had post-appendectomy IAA status classified by the pediatric surgeons, 20 patients were eliminated due to more than 20% of data for the database fields missing. From the remaining 1574 patients, 34 variables from the database were initially identified based on only data completeness (low sparsity) with further variable processing described below (Fig. 1). Variable 34, postoperative IAA, is our designated outcome variable.

Data Processing

Feature Engineering

To improve the performance of machine learning algorithms, feature engineering is necessary at times. Feature engineering is the creation of one variable from raw data or other variables to facilitate processing by the ANN. In our dataset, we engineered 4 features/variables (Fig. 1) where we derived single numerical values for age (from the date of birth), Pediatric Appendicitis Score (from individual categories), duration of operation (from operation start and end times), and count of tools used (aggregate from multiple tool categories).

Data Leakage

Data leakage may occur with any machine learning development when information from outside a training dataset is accessible to a machine-learning algorithm leading to inaccurate learning and invalid outcomes from a predictive model (eg, training and testing with the exact same dataset). Our team split data 80%/20% (training/test) before solving for missing data to avoid data leakage.[17]

Dealing With Missing Data

Some of the variables initially identified had missing values. Data imputation (Multiple Imputations by Chained Equations [MICE] algorithm) was used to replace missing data points. MICE operate under the assumption the missing data are missing at random, which means that the probability that a value is missing depends only on observed values and not on unobserved values within the variable.[18] The R-package MICE used in this project creates multiple imputations (replacement values) for

	Varaible Name	Variable Description	Numeric (Y/N)?	# of Factors	Engineered Variable?	Dervided From
1	ageOR	Age @ Operation	Y		Y	Patient DOB
2	gender	Patient Gender (Male/Female)	N	2	N	
3	raceeth	Patient Race/Ethnicity	N	4	N	
4	htcm	Hight (cm)	Y		N	
5	wtkg	Weight (Kg)	Y		N	
6	durOR(mins)	Duration of Operation (mins)	Y		Y	Operation Start & Finissh
7	sgymethod	Surgery Method	N	5	N	
8	sgyirrigation	Irrigation	N	2	N	
9	deviceCount	Total Number of Devices Used	Y		Y	Count of Tools Used
10	sgydevice___0	Stapler	N	2	N	
11	sgydevice___1	Endoloop	N	2	N	
12	sgydevice___2	Ligasure / Harmonic / EnSeal	N	2	N	
13	sgydevice___3	Hemoclips/locks	N	2	N	
14	sgydevice___4	Endocatch	N	2	N	
15	sgydevice___5	Additional ports (>3)	N	2	N	
16	sgydevice___6	Other (select to add comments)	N	2	N	
17	finaldx	Final Diagnosis (Surgeon's)	N	6	N	
18	pathdx	Pathologic Diagnosis (Pathology)	N	6	N	
19	scandaytempmaxc	Post-op Image 1 Temperature	Y		N	
20	scandaywbc	WBC @ Scan Day	Y		N	
21	antibiotics_complete	Did Patient Complete Antibiotics	N	3	N	
22	scoreCount	Pediatric Appendicitis Score	Y		Y	Count of App. Scores
23	passx___7	Pain with movement	N	2	N	
24	passx___0	RLQ pain w/ light palpation	N	3	N	
25	passx___1	Anorexia	N	4	N	
26	passx___3	Fever (>38C/101.4F)	N	5	N	
27	passx___2	Nausea/Vomitting	N	6	N	
28	passx___4	Leukocytosis (WBC >10,000)	N	7	N	
29	passx___5	Left Shift (Seg Neutrophils >75%)	N	8	N	
30	passx___6	Pain migration to RLQ	N	9	N	
31	symptomduration	Symptoms Duration	Y		N	
32	admitwbc	WBC @ Admission	Y		N	
33	postopwbc1	WBC @ Post-Operation	Y		N	
34	postopiaa	Post Operation IAA	N	2	N	

FIGURE 1. List of initial variables retrieved from research database.

Downloaded from http://journals.lww.com/aosopen by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 07/10/2023

Downloaded from http://journals.lww.com/aosopen by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 07/10/2023

multivariate missing data. The method is based on fully conditional specification, where each incomplete variable is imputed by a separate model.

Dataset Class Balancing

Our dataset was imbalanced due to the low number of IAA positive patients (246) and a higher number of negative patients (1328), resulting in an approximate 1:5 ratio. Classification using class-imbalanced data will become biased in favor of the majority class.[19] We addressed class imbalance via the popular Synthetic Minority Oversampling Technique (SMOTE) to create synthetic examples from the minority class producing class-balanced data.[20] SMOTE first selects a minority class instance (x) at random and finds its *k*-nearest minority class neighbors via the *k*-nearest neighbors' algorithm. The synthetic instance is then created by choosing one of the *k*-nearest neighbors (y) at random and connecting (x) and (y) to form a line segment in the feature space. The synthetic instances are generated as a convex combination of the two chosen instances (x) and (y). This procedure can be used to create as many synthetic examples for the minority class as are required.[21]

Feature Selection

There are multiple ways to reduce or select a feature (input variable) before training the model. In our case, for variable reduction, Random Forest methodology (with 500 decision trees) was performed. Figure 2 shows the variable importance ranking for each variable after the Random Forest. Two additional variables were identified in the Correlation Matrix (for numeric variables) resulting in 12 variables with the most significant influence on the outcome for inclusion in our ANN models. Figure 3 shows the descriptive analysis of the models' input variable data.

Predictive ANN Models

Two "reproducible" ANN with different architectures using R-package for H2O were developed to predict the risk of developing IAA post-appendectomy. For model 1, we left the classes imbalanced and used: 12 inputs, 3 hidden layers with 12 neurons each, and 1 output. For model 2, we oversampled the minority class (using SMOTE algorithm as described above) to balance both classes and used: 12 inputs, 2 hidden layers with 18 neurons each, and 1 output. Both models had the same activation function (rectified linear unit) in the hidden layers. The number of neurons and number of hidden layers for each model were chosen as generally accepted starting points when creating a new ANN. Figure 4 shows an illustration of the 2 different ANNs created for this project.

To ensure our models and methods are reproducible to provide the same results every time we run our model on a different data set, we used a set. seed () function as otherwise ANN randomly selects the weights for the input nodes. Setting the seed of R's random number generator is useful for creating simulations or random objects that can be reproduced in addition to keeping a record of all the R-libraries versions we used in our study.

RESULTS

Predictive ANN Models

Thirty-four variables were identified from the database initially based solely on data sparsity and summarized (Fig. 1), with

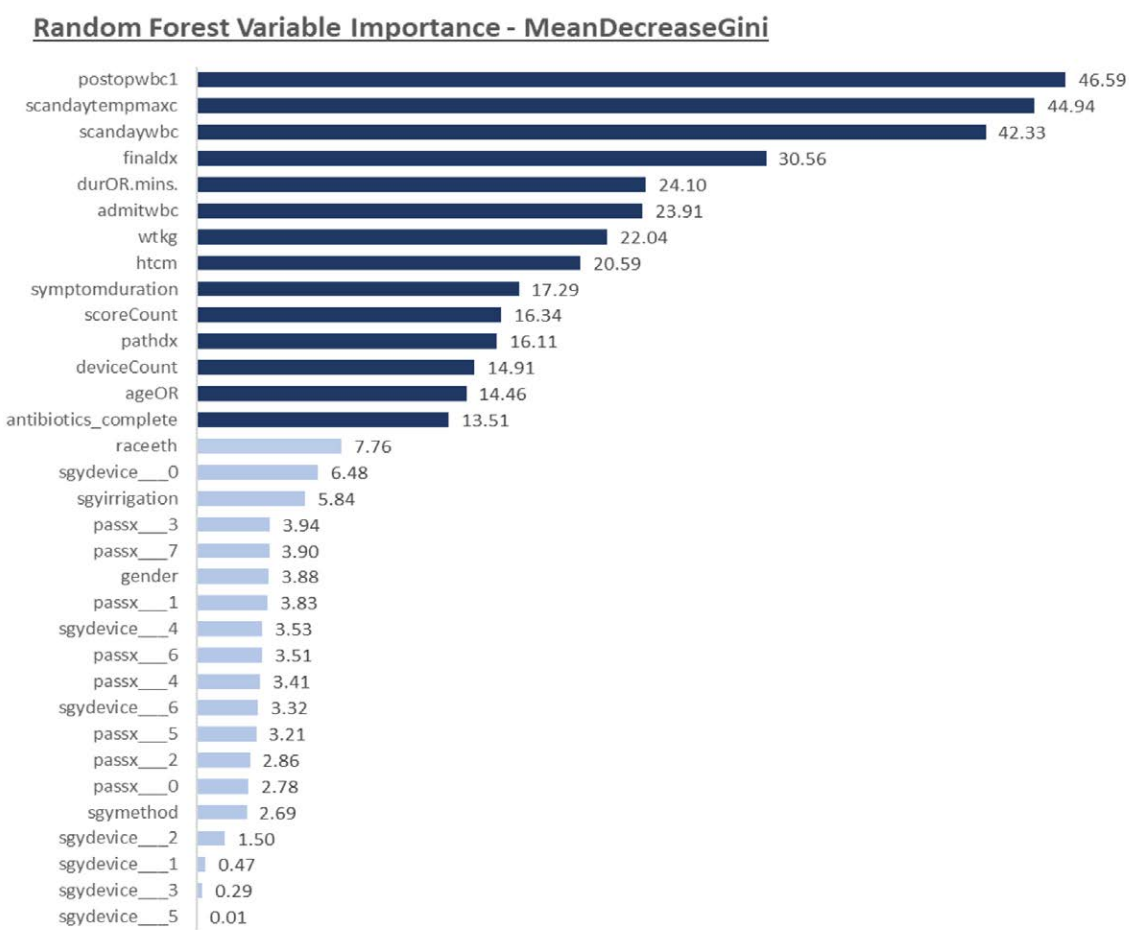

FIGURE 2. Random Forest's important variables for feature reduction.

Downloaded from http://journals.lww.com/aosopen by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 07/10/2023

1,574 patients classified by surgeons for Intra-Abdominal Abcess Post-Appendectomy for Appendicitis			
Post-IAA Status		**Negative**	**Positive**
Count		1328	246
WBC @ Admission			
	Mean	16.0	17.4
	Standard Deviation	6.3	6.6
Post-op Image Temp. (C)			
	Mean	37.50	38.04
	Standard Deviation	0.83	0.87
Height (cm)			
	Mean	144.09	139.61
	Standard Deviation	23.23	25.86
WBC @ Scan Day			
	Mean	12.1	16.6
	Standard Deviation	6.1	5.0
Weight (kg)			
	Mean	46	45
	Standard Deviation	21	23
Post Operation WBC			
	Mean	9.7	14.3
	Standard Deviation	4.2	5.5
Pediatric Appendicitis Score Count			
	Mean	5	4
	Standard Deviation	2	3
Symptoms Duration (hours)			
	Mean	42	70
	Standard Deviation	43	52
Operation Duration (min)			
	Mean	39	57
	Standard Deviation	82	24
Completed Antibiotics			
	Incomplete	117	61
	Unverified	240	64
	Complete	971	121
Pathologic Diagnosis (Pathology)			
	Simple	781	45
	Gangrenous	251	62
	Perforated	258	138
	Interval	2	0
	Normal	25	0
	Other	11	1
Final Diagnosis (Surgeon's)			
	Simple	792	28
	Gangrenous	87	5
	Perforated	436	212
	Interval	2	
	Normal	10	1
	Other	1	

FIGURE 3. ANN input variables' descriptive analysis.

variable 34 as our designated outcome variable. Via feature selection methodology as above, from these 34 variables, 12 positively impacted weighted variables were identified, incorporated, and overlapped between the 2 ANN models (Fig. 4).

After training classifying which patient had IAA post-appendectomy, model 1 achieved accuracy of 89.84%, sensitivity of 70%, and specificity of 93.61% on the test dataset, while model 2 (class-balanced via the SMOTE algorithm) achieved accuracy of 84.9%, sensitivity of 82%, and specificity of 84.6% on our test dataset.

The top 5 highly weighted variables for model 1 were postoperation white blood cells count (100%), final diagnosis by the surgeon at the time of surgery (96%), completed antibiotic course (95%), weight (89%), and operation duration in minutes (83%; Fig. 5).

For model 2, the top 5 highly weighted variables were postoperative temperature on imaging date (100%), operation

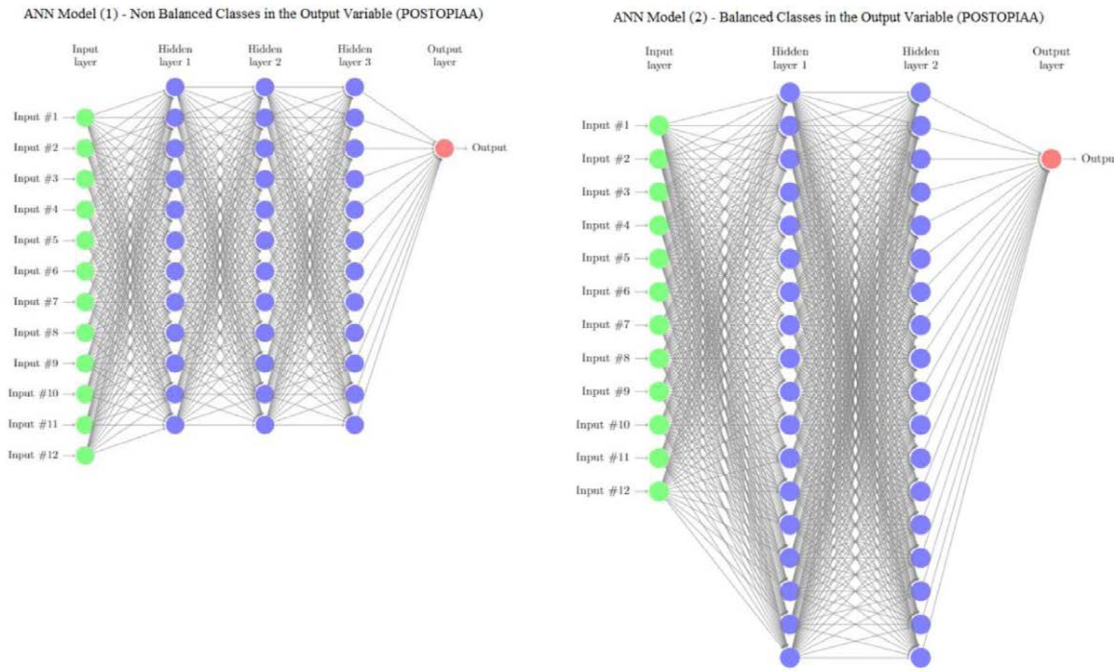

FIGURE 4. Illustrations for ANN architectures used for this project.

duration (96%), weight (94%), final diagnosis by the surgeon at the time of surgery (92%), completed antibiotic course (91%; Fig. 5).

DISCUSSION

Our results demonstrate that it is possible to develop ANN capability to predict patients who will have IAA post-appendectomy with relatively high accuracy, sensitivity, and specificity based on our dataset.

We did not identify any other application of ANN to complications of appendicitis in our literature search. There were few studies that applied ANN to detect IAA for any reason. Freed et al[22] reviewed 140 patients that underwent abdominal and pelvic computed tomography (CT scan) to detect IAA regardless of the cause and construct an ANN model based on clinical, laboratory, and radiographic examination (CT scan) and achieved sensitivity and specificity of 90% and 51%, respectively. Another study conducted by Qiu et al[23] including 263 patients used ANN compared with a logistic regression model (LRM) in predicting intra-abdominal infection (as defined as extraluminal gas in pancreatic or peripancreatic tissues on contrast CT or positive peritoneal fluid culture) in moderately severe and severe acute pancreatitis and found higher sensitivity and specificity for ANN than LRM. Both publications concluded the advantage of using ANN in predicting the intra-abdominal infection.

"Explainability" should be addressed for any newly developed artificial intelligence application model. While lacking a clear definition, explainability "can be understood as a characteristic of an artificial intelligence (AI) driven system allowing a person to reconstruct why a certain AI came up with the presented predictions."[24] There are many principles to consider for explainability of any AI methodology such as comprehensibility, fidelity, accuracy, scalability, and generality.[25] A complete discussion of explainability is beyond the scope of this article. Predictive models may have inherent explainability such as linear or LRMs. This may be in contrast to approximated explainability of "opaque" models such as ANN, whereby the weighting and pathways between the artificial neurons may not be easily interpreted in a way that humans can understand.[25] In some cases, opaque models such as ANN may perform better in predictive tasks than those with inherent explainability.[24] Amann et al[24] describe 2 levels of explainability with respect to AI-based clinical decision support tools: 1) First-level explainability seeks to describe how a system arrives at conclusions in general. 2) Second-level explainability seeks to identify features that were important for an individual prediction.[24] In the case of our (and any) multilayer ANN, we must attempt a form of post-hoc first-level explainability through exploration of feature relevance and transparency with respect to model development.[26] Second-level explainability is difficult to achieve with ANN, and yet is important for predictive models for dispute resolution between AI system and human experts, especially as attacks on the security of healthcare machine-learning algorithms have already been proposed.[24,25,27] We believe we have clearly and transparently described our ANN model development to achieve a high degree of first-level explainability.

With respect to feature relevance, highly weighted variables identified by our models support the face validity of the ANNs and contribute to first-level explainability: Several of the highly weighted variables between both models are well-known risk factors for developing IAA such as longer operation time, simple versus complicated appendicitis, body temperature, leukocytosis, and increasing weight.[28–30] As "final diagnosis by the surgeon at the time of surgery" was weighted more than 90% for model 1 and more than 95% for model 2, our ANN models do not contradict the observed experience of severity of appendicitis playing a role in likelihood of development of IAA. This particular variable incorporates surgeons' clinical judgement and experience while the ANN may reduce the impact of subjective variability by identifying interactions with additional variables.

Further exploring feature relevance, the variable "Completed antibiotic course" impacted or ANNs 95% for model 1 and 91% in model 2. In our research database, this is recorded as complete, unverified, or incomplete for either complicated or uncomplicated appendicitis. A 2005 Cochrane review concluded "Antibiotic prophylaxis is effective in the prevention of postoperative complications in appendectomized patients, whether the administration is given pre-, peri-, or postoperatively, and could

Downloaded from http://journals.lww.com/aosopen by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 07/10/2023

Downloaded from http://journals.lww.com/aosopen by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 07/10/2023

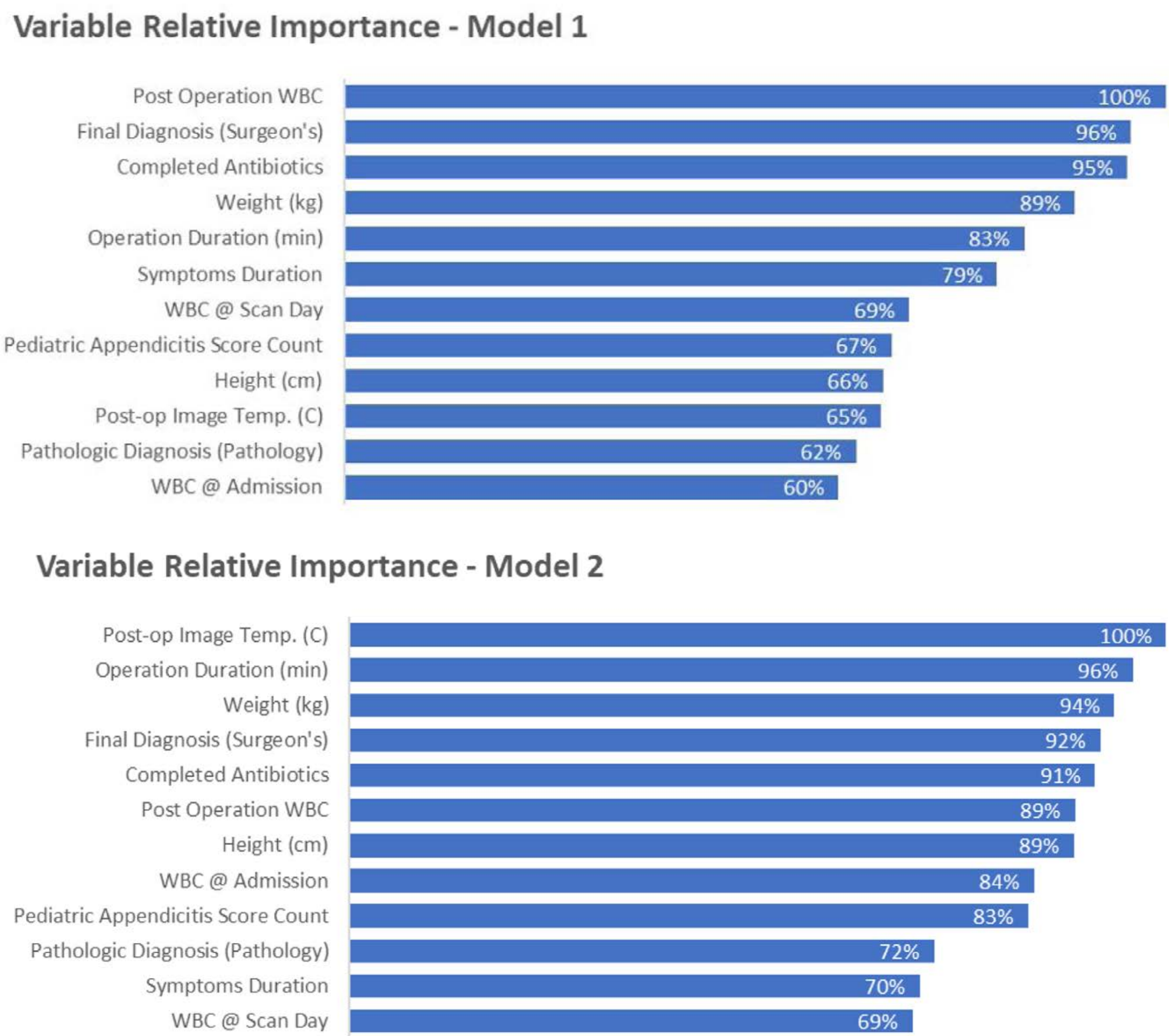

FIGURE 5. ANN input variables' relative importance.

be considered for routine in emergency appendectomies."[31] There remains variability to duration of therapy for complicated appendicitis as Surgical Infection Society guidelines state "the resolution of clinical signs of infection should be used to judge the termination point for antimicrobial therapy, antimicrobial therapy should be limited to no greater than seven days, unless it is difficult to achieve adequate source control."[32] Furthermore, there are little data regarding how postoperative antibiotics in perforated appendicitis affect outcomes. Our ANN models suggest that antibiotic course does play a role in predicting IAA, which correlates with evidence that standardized management pathways including antibiotic administration recommendations have been reported to reduce complications for both adult and pediatric appendicitis patients.[33] Walczak et al[34] developed an ANN model to predict surgical site infections from 780 surgical procedures from 9 surgical specialties. While they found that surgical antimicrobial prophylaxis bundle compliance was associated with a lower incidence of SSIs, ANN models incorporating the surgical antimicrobial prophylaxis bundle compliance as a variable had lower sensitivity but increased specificity.[34] Our ANN models suggest ensuring antibiotic administration compliance in complicated appendicitis patients should be prioritized in postoperative care.

As clinicians, it is tempting to try and study ANN predictive variables in isolation. For example with respect to surgical method, Krisher et al[35] reported that for patients with perforated appendicitis, laparoscopic appendectomy had a risk ratio of 5.6 for postoperative IAA when compared with patients who had perforated appendicitis and open appendectomies. A Cochrane review of laparoscopic versus open surgical method for suspected appendicitis also concluded that laparoscopic procedures appeared to be associated with higher risk of IAA, though the effect was not identified in pediatrics.[36] Other meta-analyses contradict the findings from the Cochrane review.[37,38] In many instances, perforated appendicitis is not identified until the surgical procedure has started, logistically complicating selecting surgical modality based on a surgical diagnosis. Laparoscopy has become the primary modality of appendectomy, including at our own institution where 96.6% overall and 94% of the 648 patients with perforated appendicitis had a laparoscopic appendectomy. Accordingly, aligning with first-level explainability for our model development, surgical method was found to have little influence on postoperative IAA prediction in the first feature selection step of Random Forest analysis. Examining our own and Krisher et al's[35] data from another perspective, 72% of Krisher et al's[35] patients and 86% of our own patients that

Downloaded from http://journals.lww.com/aosopen by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 07/10/2023

developed post-appendectomy IAA had perforated appendicitis. In our ANN models, the variable "Final Diagnosis (Surgeon's) - final diagnosis by the surgeon at the time of surgery" (ie, simple, gangrenous, perforated) was a highly predictive variable rather than the surgical method. The true benefit of ANN lies in the nonlinear analysis of multiple variables' relationships. In our example, when all 12 final variables are taken together, it paints a picture of a patient with more complicated appendicitis, strengthening first-level explainability.

Additionally, prehospital factors such as insurance status or limited English proficiency may predict the probability of perforated appendicitis.[39,40] The database we used for our models did not include primary language spoken as a variable. Insurance status was included in the database, but due to missing data for the majority of patients, this variable was not included in development of the ANNs. Since these prehospital factors may be either unknown or out of the surgeon's control, ANN models can identify patients likely to develop an IAA post-appendectomy who may not be immediately apparent to the surgeon. When these patients are identified via ANN, the clinician can focus on clinical pathway adherence and correct antibiotic prescribing. These ANNs are also novel in that they incorporate patient and clinical practice factors with clinician judgement.

This study's primary limitation is that it is a single-center study. More sophisticated data processing techniques were required due to data sparsity and class imbalance in our output variable reducing explainability. Having access to larger dataset from different locations may help reduce data sparsity and class imbalance, while potentially identifying additional variables of importance (such as prehospital factors if available in the database) and additional hidden layers and interactions between nodes. Caution should be exercised when incorporating new variables to avoid what is known as the "Clever Hans" phenomenon whereby meta-data may drive the prediction rather than the actual data itself.[41] Additional variables may also reduce first- and second-level explainability.[25] Prospective external validation on additional patients' needs to be performed. Generalizability to other institutions requires external validation as well. Another limitation is that not all ANN architectures were explored. Finally, we did not compare our ANN model to experienced clinicians. It is plausible that an experienced surgeon given the same dataset may be more sensitive and specific, though we speculate time to complete the task would likely be significantly more. We would also like to explore incorporation of the ANN into our institutions clinical pathway guideline and determine impact on clinical outcomes.

In conclusion, ANN applied to selected variables can accurately predict patients who will have IAA post-appendectomy, a novel application of ANN. Both models are reproducible with the same 12 input variables, achieved high accuracy, sensitivity, and specificity. Our ANNs are a state-of-the-art method to incorporate patient, surgeon, and clinical practice variables for identifying avenues for optimizing post-appendectomy care.

REFERENCES

1. Anandalwar SP, Cameron DB, Graham DA, et al. Association of intraoperative findings with outcomes and resource use in children with complicated appendicitis. JAMA Surg. 2018;153:1021–1027.
2. Diefenbach K, Lawrence Moss R. Intra-abdominal, visceral, and retroperitoneal abscesses. In: Long S, ed. Principles and Practice of Pediatric Infectious Diseases. Elsevier; 2018:432–436.
3. Ponsky TA, Hafi M, Heiss K, et al. Interobserver variation in the assessment of appendiceal perforation. J Laparoendosc Adv Surg Tech A. 2009;19(suppl 1):S15–S18.
4. St Peter SD, Sharp SW, Holcomb GW III, et al. An evidence-based definition for perforated appendicitis derived from a prospective randomized trial. J Pediatr Surg. 2008;43:2242–2245.
5. Hong WD, Ji YF, Wang D, et al. Use of artificial neural network to predict esophageal varices in patients with HBV related cirrhosis. Hepat Mon. 2011;11:544–547.
6. Delibegovic S, Brkic S, Nuhbegovic S. Possibility of simplification of APACHE II scoring system in the prediction of the outcome in critically ill patients with perforative peritonitis. Med Arh. 2009;63:249–251.
7. Chen J, Pan QS, Hong WD, et al. Use of an artificial neural network to predict risk factors of nosocomial infection in lung cancer patients. Asian Pac J Cancer Prev. 2014;15:5349–5353.
8. Chen J, Chen J, Ding HY, et al. Use of an artificial neural network to construct a model of predicting deep fungal infection in lung cancer patients. Asian Pac J Cancer Prev. 2015;16:5095–5099.
9. Jha S, Topol EJ. Adapting to artificial intelligence: radiologists and pathologists as information specialists. JAMA. 2016;316:2353–2354.
10. Cheng CA, Chiu HW. An artificial neural network model for the evaluation of carotid artery stenting prognosis using a national-wide database. Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Int Conf. 2017;2017:2566–2569.
11. Shahid N, Rappon T, Berta W. Applications of artificial neural networks in health care organizational decision-making: a scoping review. PLoS One. 2019;14:e0212356.
12. Ahmed FE. Artificial neural networks for diagnosis and survival prediction in colon cancer. Mol Cancer. 2005;4:29.
13. Zhang L, Tan J, Han D, et al. From machine learning to deep learning: progress in machine intelligence for rational drug discovery. Drug Discov Today. 2017;22:1680–1685.
14. Hsu H. How Do neural network systems work? Computer History Museum Blog. 2020. Available at: https://computerhistory.org/blog/how-do-neural-network-systems-work/. Accessed March 11, 2021.
15. Hardesty L. Explained: neural networks. Massachussetts Institute of Technology News. 2017. Available at: https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414. Accessed March 11, 2021.
16. Bruny JL, Hall BL, Barnhart DC, et al. American College of Surgeons National Surgical Quality Improvement Program Pediatric: a beta phase report. J Pediatr Surg. 2013;48:74–80.
17. Brownlee J. Data Leakage in Machine Learning. Machine Learning Mastery. 2020. Available at: https://machinelearningmastery.com/data-leakage-machine-learning/. Accessed March 11, 2021.
18. Azur MJ, Stuart EA, Frangakis C, et al. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res. 2011;20:40–49.
19. Brownlee J. A gentle introduction to imbalanced classification. Machine Learning Mastery. 2020. Available at: https://machinelearningmastery.com/what-is-imbalanced-classification/. Accessed March 11, 2021.
20. Hoens TR, Chawla N. Imbalanced datasets: from sampling to classifiers. In: He H, Ma Y, eds. Imbalanced Learning: Foundations, Algorithms, and Applications. Wiley; 2013:44–47.
21. Blagus R, Lusa L. SMOTE for high-dimensional class-imbalanced data. BMC Bioinformatics. 2013;14:106.
22. Freed KS, Lo JY, Baker JA, et al. Predictive model for the diagnosis of intraabdominal abscess. Acad Radiol. 1998;5:473–479.
23. Qiu Q, Nian YJ, Tang L, et al. Artificial neural networks accurately predict intra-abdominal infection in moderately severe and severe acute pancreatitis. J Dig Dis. 2019;20:486–494.
24. Amann J, Blasimme A, Vayena E, et al; Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak. 2020;20:310.
25. Belle V, Papantonis I. Principles and practice of explainable machine learning. Front Big Data. 2021;4:688969.
26. Cutillo CM, Sharma KR, Foschini L, et al; MI in Healthcare Workshop Working Group. Machine intelligence in healthcare-perspectives on trustworthiness, explainability, usability, and transparency. NPJ Digit Med. 2020;3:47.
27. Mozaffari-Kermani M, Sur-Kolay S, Raghunathan A, et al. Systematic poisoning attacks on and defenses for machine learning in healthcare. IEEE J Biomed Health Inform. 2015;19:1893–1905.
28. Schlottmann F, Sadava EE, Peña ME, et al. Laparoscopic appendectomy: risk factors for postoperative intraabdominal abscess. World J Surg. 2017;41:1254–1258.
29. Davies DA, Yanchar NL. Appendicitis in the obese child. J Pediatr Surg. 2007;42:857–861.
30. Reid RI, Dobbs BR, Frizelle FA. Risk factors for post-appendicectomy intra-abdominal abscess. Aust N Z J Surg. 1999;69:373–374.
31. Andersen BR, Kallehave FL, Andersen HK. Antibiotics versus placebo for prevention of postoperative infection after appendicectomy. Cochrane Database Syst Rev. 2005:CD001439.
32. Mazuski JE, Tessier JM, May AK, et al. The surgical infection society revised guidelines on the management of intra-abdominal infection. Surg Infect (Larchmt). 2017;18:1–76.

33. Helmer KS, Robinson EK, Lally KP, et al. Standardized patient care guidelines reduce infectious morbidity in appendectomy patients. Am J Surg. 2002;183:608–613.
34. Walczak S, Davila M, Velanovich V. Prophylactic antibiotic bundle compliance and surgical site infections: an artificial neural network analysis. Patient Saf Surg. 2019;13:41.
35. Krisher SL, Browne A, Dibbins A, et al. Intra-abdominal abscess after laparoscopic appendectomy for perforated appendicitis. Arch Surg. 2001;136:438–441.
36. Jaschinski T, Mosch CG, Eikermann M, et al. Laparoscopic versus open surgery for suspected appendicitis. Cochrane Database Syst Rev. 2018;11:CD001546.
37. Liu Z, Zhang P, Ma Y, et al. Laparoscopy or not: a meta-analysis of the surgical effects of laparoscopic versus open appendicectomy. Surg Laparosc Endosc Percutan Tech. 2010;20:362–370.
38. Ohtani H, Tamamori Y, Arimoto Y, et al. Meta-analysis of the results of randomized controlled trials that compared laparoscopic and open surgery for acute appendicitis. J Gastrointest Surg. 2012;16:1929–1939.
39. Smink DS, Fishman SJ, Kleinman K, et al. Effects of race, insurance status, and hospital volume on perforated appendicitis in children. Pediatrics. 2005;115:920–925.
40. Levas MN, Dayan PS, Mittal MK, et al; Pediatric Emergency Medicine Collaborative Research Committee of the American Academy of Pediatrics. Effect of Hispanic ethnicity and language barriers on appendiceal perforation rates and imaging in children. J Pediatr. 2014;164:1286–91.e2.
41. Lapuschkin S, Wäldchen S, Binder A, et al. Unmasking Clever Hans predictors and assessing what machines really learn. Nat Commun. 2019;10:1096.

Downloaded from http://journals.lww.com/aosopen by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IIQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 07/10/2023