

A Novel Application of Neural Networks to Identify Potentially Effective Combinations of Biologic Factors for Enhancement of Bone Fusion/Repair

Neural Networks for Bone Fusion

Albert T Anastasio	Department of Orthopaedic Surgery Duke University Medical Center, Durham, NC 27710
Bailey S Zinger	Chemical and Biological Engineering Department University of Colorado at Boulder, Boulder, CO 80302
Thomas J Anastasio	Department of Molecular and Integrative Physiology University of Illinois at Urbana-Champaign, Urbana, IL 61801

address for correspondence:

Albert T Anastasio, MD
Duke University Hospital
2301 Erwin Rd, Durham, NC 27710
(217)-418-9132

ABSTRACT

Introduction: The use of biologic adjuvants (orthobiologics) is becoming commonplace in orthopaedic surgery. Amongst other applications, biologics are often added to enhance fusion rates in spinal surgery and to promote bone healing in complex fracture patterns. Generally, orthopaedic surgeons use only one biomolecular agent (ie allograft with embedded bone morphogenetic protein-2) rather than several agents acting in concert. Bone fusion, however, is a highly multifactorial process and it likely could be more effectively enhanced using biologic factors in combination, acting synergistically. We used artificial neural networks to identify combinations of orthobiologic factors that potentially would be more effective than single agents.

Methods: Available data on the outcomes associated with various orthopaedic biologic agents, electrical stimulation, and pulsed ultrasound were curated from the literature and assembled into a form suitable for machine learning. The best among many different types of neural networks was chosen for its ability to generalize over this dataset, and that network was used to make predictions concerning the expected efficacy of 2400 medically feasible combinations of 9 different agents and treatments.

Results: The most effective combinations were high in the bone-morphogenetic proteins (BMP) 2 and 7 (BMP2, 15mg; BMP7, 5mg), and in osteogenin (150ug). In some of the most effective combinations, electrical stimulation could substitute for osteogenin. Some other effective combinations also included bone marrow aspirate concentrate. BMP2 and BMP7 appear to have the strongest pairwise linkage of the factors analyzed in this study.

Conclusions: Artificial neural networks are powerful forms of artificial intelligence that can be applied readily in the orthopaedic domain, but neural network predictions improve along with the amount of data available to train them. This study provides a starting point from which networks trained on future, expanded datasets can be developed. Yet even this initial model makes specific predictions concerning potentially effective combinatorial therapeutics that should be verified experimentally. Furthermore, our analysis provides an avenue for further research into the basic science of bone healing by demonstrating agents that appear to be linked in function.

CLINICAL RELEVANCE

Bone healing is a highly multifactorial process, and it likely could be more effectively enhanced using combinations of factors rather than single factors in isolation. This study provides a starting point for an integration of biomedical experimentation and computational AI that ultimately could lead to highly sophisticated combinatorial treatments for bone repair and other applications in orthopaedic medicine.

KEYWORDS

Orthopaedic surgery; bone fusion; bone repair; BMP; BMAC; PRP; osteogenin; biologic therapy; artificial intelligence; machine learning; deep neural networks; combinatorial therapy; polypharmacy

INTRODUCTION

Bone repair is a highly multifactorial process involving a wide array of molecular and cellular factors.[1] Orthopaedic surgeons have manipulated these factors by administering various biologic agents in order to augment bone repair.[2] In most cases, surgeons have administered only one biologic agent. Considering the physiological complexity of the process, it is reasonable to suggest that superior bone repair could be achieved using biologic factors in combination.

Combinatorial explosion prohibits exhaustive experimental evaluation of the full set of possible combinations. An alternative is to use computational methods to extrapolate, or *generalize*, from existing data and predict which combinations would be the most effective, and then expend experimental resources to evaluate only those. The most powerful AIs in use today for making accurate predictions are artificial neural networks (ANNs).[3, 4] ANNs are composed of many, highly interconnected neuron-like elements known as units, which can be arranged in layers or circuits.

ANNs are computational devices that process information from their input units to produce a pattern of activation at their output units. Feedforward networks have their units arranged in layers. The simplest feedforward networks have only two layers of units: input and output. More complex feedforward networks have one or more layers of hidden units, so called because they are interposed between the input and output layers. Feedforward ANNs are considered deep if they have more than two hidden layers.[5] Recurrent networks have their units arranged in circuits. Multiple processing layers, or circuits, are needed when the production of useful output patterns requires the processing of complex interactions among the inputs.

ANNs are trained via machine learning on a set of input/desired-output examples. They have been applied in many domains of biomedicine.[6-8] The most extensive medical applications of ANNs have been in radiology.[9-12] Generally in these applications the inputs are the pixels of (usually MRI) images, and the desired outputs are the components of known radiological diagnoses. Once trained, the ANN could generalize from its training data and could make a diagnosis from an image on which it has not been trained. It is likely that clinicians will soon use ANNs adjunctively in radiological diagnosis.

Other applications under development of ANNs in medicine include cancer diagnosis from gene expression data[13, 14], heart-disease diagnosis from electrocardiogram data[15], osteoporosis diagnosis from bone-density data[16], and diabetes diagnosis from blood chemistry data[17, 18]. More recent applications involve patient records as inputs.[19-21] ANNs have been applied to orthopaedic patient records, to predict outcomes such as bone fracture healing or mortality following hip fracture.[22, 23]

The application of various biologic agents to bone repair is a rapidly growing subfield of orthopedics. Over the past few decades, many reports have demonstrated the benefits of specific agents on post-surgical bone fusion rates. Still, to our knowledge, AI has not been applied in this domain.

The usefulness of an ANN derives from its ability to *generalize* beyond its training data, so that it can predict the correct output for inputs on which it has not been trained. We collected a large amount of the available data, organized it into a form suitable for machine learning, and used it to train an ANN with an architecture that we had determined beforehand would generalize well over the dataset. We used this ANN to explore potential combinatorial therapies within the realm of orthobiologic adjuncts.

METHODS

Our study design consisted of six steps: (1) assemble a dataset on the outcomes associated with the use of various agents and organize them into a form suitable for machine learning; (2) build a series of ANNs with increasingly complex architectures and processing potentials; (3) determine for each network type its optimal machine learning parameters; (4) assess the ability of each network type to generalize over the dataset; (5) train the best generalizing network on the dataset and use it to predict the efficacy of a large set of combinations of the factors on which it had been trained; and (6) analyze the predictions and determine which combinations of factors are potentially the most effective, and should be experimentally verified. The first five steps are methodological and are summarized here in the Methods section. Further methodological details are available in Supplementary Texts S1 – S4. Step (6) is elaborated in the Results section.

Step 1. We curated the dataset from the experimental literature and organized the data into input/desired-output pairs. In total, 17 factors (active orthobiologic agents and their vehicles of administration, or other nonpharmacological treatment types) constituted the inputs, and 26 outcomes (metrics quantifying the efficacy for improvement of bone healing due to the agents) constituted the desired outputs. The input/desired outputs are quantified precisely in the dataset according to their units as appropriate. We used as many factors as inputs, and outcomes as desired outputs, as were available in the literature, in order to maximize the amount of ANN training data.

Step 2. We evaluated the generalizability of 16 different ANN types. Our 8 basic types were feedforward with 0, 1, 2, 3, 5, 7, or 10 hidden layers, and a recurrent network with a hidden circuit. All hidden layers (and the hidden circuit) were composed of 100 units. We evaluated each of these 8 network types with and without an autoencoder layer. An autoencoder is the hidden-unit representation developed by a network that learns to reproduce its own input at the output. Placing an autoencoder layer after the input layer can improve the generalizability of an ANN that is trained to produce a desired output for every input.

Steps 3 and 4. We optimized the parameters of the machine-learning algorithm used to train each of the network types, and then tested the ability of each type to generalize over the dataset following training. We found that the feedforward ANN composed of an input layer, an autoencoder layer, 2 hidden layers, and an output layer exhibited the best generalizability. A diagram of this ANN is shown in Figure 1.

Step 5. Following ANN training, we used clinical judgement in deciding which combinations of factors to evaluate, and which outcomes to use in assessing the predicted post-surgical benefit of those selected combinations. We generated a set of 2400 combinations of 9 of the factors that were included among the 17 inputs in the dataset. These factors were chosen because they could be combined appropriately in a surgical setting. The 9 chosen agents were bone-morphogenic protein-2 (BMP2), bone-morphogenic protein-7 (BMP7), osteogenin (OG), platelet-derived growth factor (PDGF), bone marrow aspirate concentrate (BMAC), and platelet rich plasma (PRP).[2] The vehicles carrying these agents varied greatly among published studies, so we included the most common one, exogenous bone graft (EBG), as a stand-in for all vehicles. Pulsed ultrasound (US) and electrical stimulation (ES) were also chosen for the combination screen because they have been shown to increase bone healing rates.[24, 25]

We quantized input levels in order to generate a finite number of input combinations. For the combination screen the factor BMP7 takes 2 levels; BMP2, OG, and PU each take 4 levels; and PDGF takes 5 levels in their ranges. The factors ES, BMAC, and PRP are either present or absent. EBG, as the common vehicle of administration, is present in all combinations.

We further constrained the number of combinations for reasons of practicality. Although PU and ES could theoretically be used in combination, the feasibility of carrying out this dual therapy in practice is low. Due to the additional operative time of harvesting PRP and BMAC to use at the bone healing site, we decided to not to include combinations involving both agents. Editing according to these constraints left 2400 combinations.

To screen the combinations for efficacy, we set the input in turn to each one of the 2400 combinations: the 9 input units corresponding to BMP2, BMP7, OG, PDGF, PU, ES, BMAC, PRP, and EBG took their values as specified for that combination; the other 8 of the 17 input units took value 0. We then computed the activities in response to each input of the 26 output units. Due to randomness inherent in the machine-learning algorithms, ANNs of the same type trained on the same dataset can nevertheless vary slightly. Therefore, the best predictions are derived from the averaged outputs of several ANNs.[26] We based our predictions on the averaged outputs of 10 separately trained ANNs of the type shown in Figure 1.

To compute a relative efficacy measure for each factor combination, we combined 17 of the 26 averaged output unit activations into a single number. We chose these 17 outcomes because they best assessed the degree of bone healing and functional outcome across studies. The chosen outcomes are distraction rate (DR), bone formation at 3 months (BF3), bone formation at 6 months (BF6), mineralized tissue volume/total tissue volume (MV/TV), 1-level posterior lumbar fusion rate (PLF-FR), Oswestry disability index improvement (ODI), fusion rate (FR), fracture healing percentage (FH), Oswestry Score (OW), radiographic outcome (RO), histomorphometric outcome (HO), implant survival percentage (IS), time to achieve full weight bearing/clinical healing (TWB/CH), mean time to radiographic union (TRU), need for repeat bone grafting (RBG), not healed at end of trial (NH), and need for dynamization (DY). We flipped the outputs whose high score indicated poor efficacy, normalized all outputs into the range [0, 1] and then averaged the 17 normalized outputs to arrive at a single-number efficacy score. By this relative measure, perfectly effective and ineffective combinations would have efficacy scores of 1 and 0, respectively.

RESULTS

We trained the ANN with the best generalizability (Figure 1) to achieve a good but not perfect match between its actual and desired outputs, because the overtraining required to achieve a perfect match would impair its ability to generalize. Comparison of the desired and actual output images for an example ANN (Figure 2) shows that the agreement is good but not perfect. Precisely this sort of relationship would be expected for an ANN that could generalize beyond its training data.

We rank-ordered the predicted efficacy scores for the 2400 combinations (Figure 3). They ranged from about 0.30 to almost 0.75 and so covered almost half of the possible [0, 1] range. The efficacy scores seemed to plateau for the most effective several hundred combinations.

The 2400 rank-ordered combinations are shown in two separate images in Figure 4: one for all 2400 combinations and another for the top 200. Analysis of the top 200 reveals some statistically significant, pairwise correlations among the 9 factors (Table 1). BMP2 and BMP7 are positively correlated, while OG and PDGF are negatively correlated. PDGF is negatively or positively correlated with BMAC or PRP, respectively. PU and ES, and likewise BMAC and PRP, are also negatively correlated, but this is due largely to constraints in the design of the combination screen (see Methods).

The 10 best factor combinations show some consistent similarities, and some consistent differences with the 10 worst combinations. BMP2 and BMP7 are at their highest levels in the 10 best combinations, while they are 0 in the 10 worst combinations, and this is consistent with the positive correlation observed between BMP2 and BMP7 in the top 200 combinations. OG tends to be at its highest levels in the 10 best combinations but is 0 in the 10 worst combinations. In contrast, PDGF tends to be at its lowest or highest levels in the 10 best or worst combinations, respectively, and this is consistent with the negative correlation observed between OG and PDGF in the top 200 combinations.

The analysis suggests that the most effective combinations are high in BMP2, BMP7, and OG, but low in PDGF (see Table 2 and its caption for quantification of amounts). These 4 factors seem to be the most determinative of the best factor combinations. BMAC appears in some of the 10 best combinations but in none of the 10 worst, and this is consistent with the negative correlation between PDGF and BMAC. PRP is absent from all 10 best and worst combinations. The 10 best and 10 worst combinations seem indifferent to the levels of PU and ES, with the potentially important exception that ES appears in some of the 10 best combinations that lack OG.

The 2400 combinations in the screen include the null combination (ie none of the 9 factors are present except for EGB, the common vehicle), and all combinations in which EGB and 1 other factor only are present. The analysis clearly indicates that combinations of several orthobiologic factors would be more effective than any single factor alone. The analysis indicates that combinations of BMP2, BMP7, and OG, perhaps including ES or BMAC, each at the high end of their ranges as reported in relevant studies, should outperform combinations that lack those components. Experimental verification of these predictions could lead to the development of orthobiologic factor combinations that outperform single factors for the enhancement of bone repair.

DISCUSSION

To properly situate our model within the orthopaedic literature, it is necessary to distinguish between process-driven and data-driven models. Process-driven models represent processes explicitly. There is a long tradition of process-driven modeling in bone fracture healing (see [1] for review). Process-driven models are valuable in that they explicitly describe the processes involved, but they are limited to what is known about the processes themselves. This limits their predictive power.

Data-driven models are built almost entirely on observed input-output relationships, without regard for the specifics of the underlying processes. Data-driven models offer little mechanistic insight, but they provide a powerful means to leverage all available data for predicting the outputs to novel inputs. Deep neural networks are the premier form of data-driven modeling in AI today. The multilayered ANN we chose to make our predictions (Figure 1) is a deep neural network.[5] To our knowledge, our model is the first data-driven, deep neural network model of the relationship between biologic factors and bone repair.

Even though it is data-driven, our model may indicate avenues for further research into the molecular physiology of bone healing. For example, OG (osteogenin, or bone morphogenic protein-3 (BMP3)) and ES seem to act interchangeably in our model. Interestingly, ES has been shown to upregulate BMPs 2 through 8, and is effective in upregulating BMP3 (also called OG) in cultured bone cells.[27] The fact that our model is able to postdict previously known molecular pathways advocates for its use in predicting previously unknown molecular interactions.

The main limitation in our study was in the size and composition of the dataset on which we trained our ANN. At 225 input/desired-output training patterns, our dataset is large in comparison with other datasets that are curated from the literature, but still very small in comparison with datasets used to train ANNs in many applications. Also, most of the input/desired-output patterns in our dataset included only one active orthobiologic factor. The risk in training mainly on single factors is that the network would fail to learn interactions among them but in our case, it seems that this did not occur.

If machine learning failed to pick up interactions, then the simplest ANN, that composed only of input and output layers, would have generalized as well as, if not better than, ANNs with hidden layers (or circuits) intervening between input and output (see Supplemental Text S4). The fact that the ANN that generalized best over our dataset was a multilayered network strongly suggests that it did learn some of the interactions between the factors.

The best way to remedy the main limitation in this analysis is to train deep neural networks on larger datasets containing more combinations of factors. The analysis already suggests both good and bad combinations that could be explored experimentally. Any and all new data on the outcomes for bone healing associated with orthobiologic factors administered alone, or better, in combination could be added to the training dataset and would improve the ability of an ANN to identify combinations of factors with the potential to outperform single agents.

ACKNOWLEDGEMENTS

This work was conducted in the absence of any public or private funding.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

TABLES

Significant Pairwise Correlations among Factors in Top 200 Orthobiologics Combinations			
factor 1	factor 2	correlation coefficient	p value
BMP2	BMP7	0.1962	0.0054
OG	PDGF	-0.2538	0.0003
PDGF	BMAC	-0.2291	0.0011
PDGF	PRP	0.2265	0.0013
PU	ES	-0.5444	0.0000
BMAC	PRP	-0.3690	0.0000

Table 1. Statistically significant pairwise correlations among the factors in the top 200 orthobiologics combinations as determined from the neural network. BMP2 and BMP7 are positively correlated, while OG and PDGF are negatively correlated. PDGF is negatively or positively correlated with BMAC or PRP, respectively. PU and ES are negatively correlated, and BMAC and PRP are also negatively correlated. The very low (but still > 0) p values associated with those last two negative correlations is attributed mainly to constraints in the design of the combination screen by which PU and ES could not be nonzero together, and BMAC and PRP likewise could not be nonzero together. All values including p values are reported to four significant places.

Ten Best Orthobiologics Combinations									
rank	BMP2	BMP7	OG	PDGF	PU	ES	BMAC	PRP	EBG
1	15	5	150	0	150	0	1	0	1
2	15	5	150	0	100	0	1	0	1
3	15	5	150	0	150	0	0	0	1
4	15	5	0	0	0	1	0	0	1
5	15	5	150	50	150	0	0	0	1
6	15	5	150	0	100	0	0	0	1
7	15	5	150	0	50	0	1	0	1
8	15	5	150	50	150	0	1	0	1
9	15	5	0	0	0	0	0	0	1
10	15	5	0	50	0	1	0	0	1
Ten Worst Orthobiologics Combinations									
rank	BMP2	BMP7	OG	PDGF	PU	ES	BMAC	PRP	EBG
2391	0	0	0	100	100	0	0	0	1
2392	0	0	0	100	50	0	0	0	1
2393	0	0	0	150	150	0	0	0	1
2394	0	0	0	150	100	0	0	0	1
2395	0	0	0	150	0	0	0	0	1
2396	0	0	0	200	150	0	0	0	1
2397	0	0	0	150	50	0	0	0	1
2398	0	0	0	200	100	0	0	0	1
2399	0	0	0	200	50	0	0	0	1
2400	0	0	0	200	0	0	0	0	1

Table 2. The 10 best and 10 worst combinations of orthobiologic factors, as determined from the neural network. Combinations are ranked, best to worst, out of the total of 2400 combinations in the computational screen. BMP2 and BMP7 are at their highest levels in the 10 best combinations, while they are 0 in the 10 worst combinations. OG tends to be at its highest or lowest levels in the 10 best or worst, respectively, while PDGF tends in the opposite direction. The 10 best and worst combinations seem indifferent to the level of PU. ES or BMAC is present in some of the 10 best but in none of the 10 worst. PRP is absent from all 10 best and worst combinations. EGB is present in them all but it is present in all 2400 combinations by design and is included only for completeness. Units: BMP2 and BMP7 are in milligrams, OG and PDGF are in micrograms, and PU is in total treatment days; the other inputs are either present or absent. The levels (dosages, intensities, amounts, etc) of all inputs are in the ranges as reported in published studies.

FIGURES

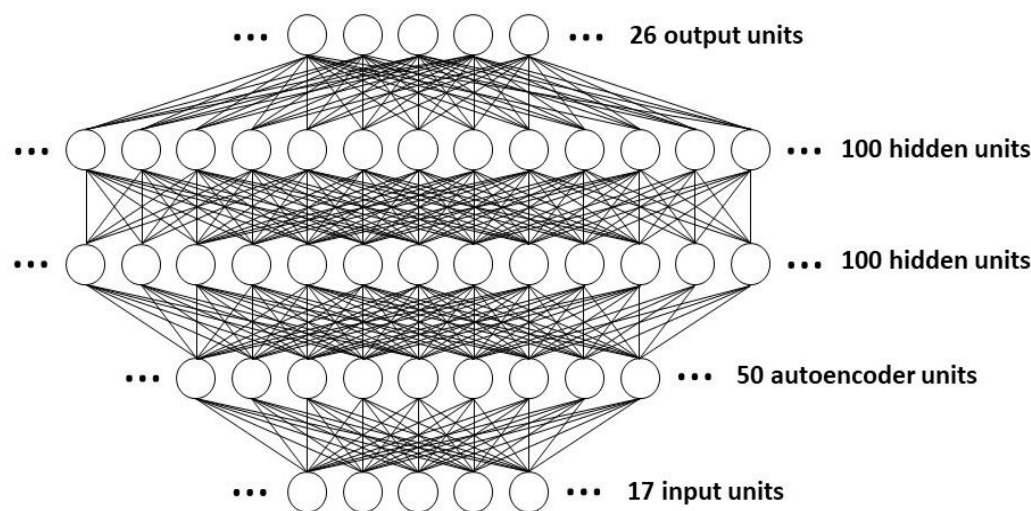


Figure 1. Diagram of the artificial neural network used to predict the efficacy of 2400 combinations of orthobiologic factors. The feedforward network has two layers of hidden units. The first hidden layer receives the input not directly but only after it has been processed by a layer of autoencoder units. The 17 input units project to a layer of 50 autoencoder units, which project to the first layer of 100 hidden units, which in turn project to the second layer of 100 hidden units, which finally project to the set of 26 output units. The connectivity between layers is complete in that each unit in a previous layer projects to every unit in a subsequent layer. The weights from the input layer to the autoencoder layer are trained separately, and then held fixed while the other weights in the network are trained.

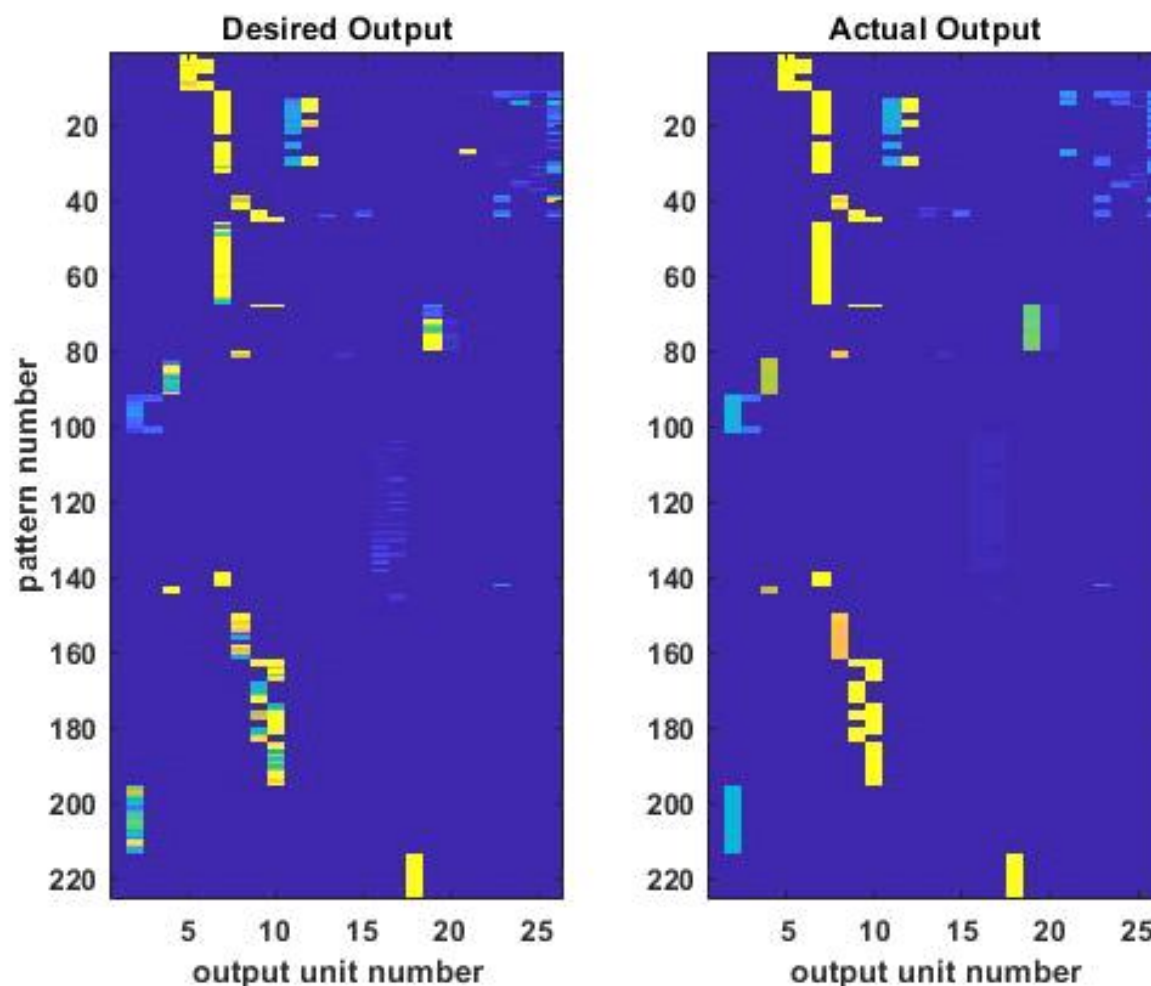


Figure 2. Desired and actual outputs of the neural network after training. The values of the 26 output units for each of the 225 input/desired-output patterns in the training dataset are shown as images, separately for the desired (left) and actual (right) outputs. The unscaled values of the outputs range from 2 (deep blue) to 277 (bright yellow). Close inspection reveals that the match between the desired and actual outputs is good but not perfect. This is expected due to ambiguity in the training dataset, which is derived directly from the experimental data of multiple labs that often reported different outputs for the same inputs. The pattern of agreement in general, with disagreement in detail, indicates that the neural network has learned to generalize from the data in the training dataset. The ability to generalize is central to the ability of the neural network predict the outputs for combinations of inputs on which it has not been trained.

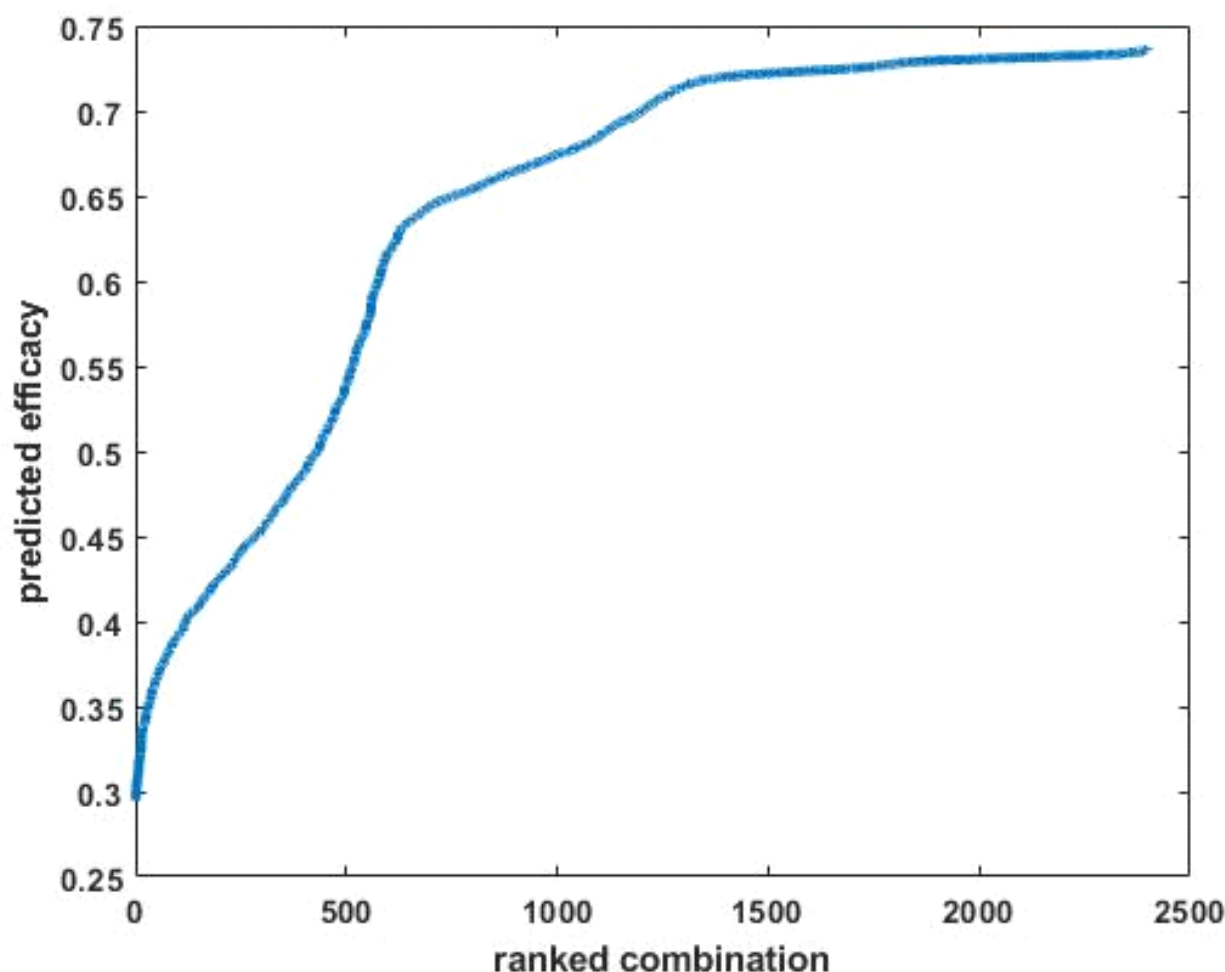


Figure 3. The efficacies of 2400 combinations of orthobiologic factors as predicted by the neural network. The scores are sorted from least to most effective. Effective combinations were defined as those that were high in 12 / 26 outputs (DR, BF3, BF6, MV/TV, PLF-FR, ODI, FR, FH, OW, RO, HO, and IS), and low in 5 / 26 outputs (TWC/CH, TRU, RBG, NH, and DY). The remaining 9 / 26 outputs were not included in the efficacy measure (see text). For the purposes of ranking, the outputs that should be low were flipped, all outputs were scaled in the range [0, 1], and the output values were averaged. By this measure, which is relative to the maximal and minimal output values, the highest possible efficacy of 1 would be obtained if all of the outputs that should be high / low were at their maximal / minimal levels for that output, and vice-versa for the lowest possible efficacy of 0. The predicted efficacies for the 2400 combinations in the screen varied widely over the [0, 1] range and nearly plateaued for the most effective several hundred.

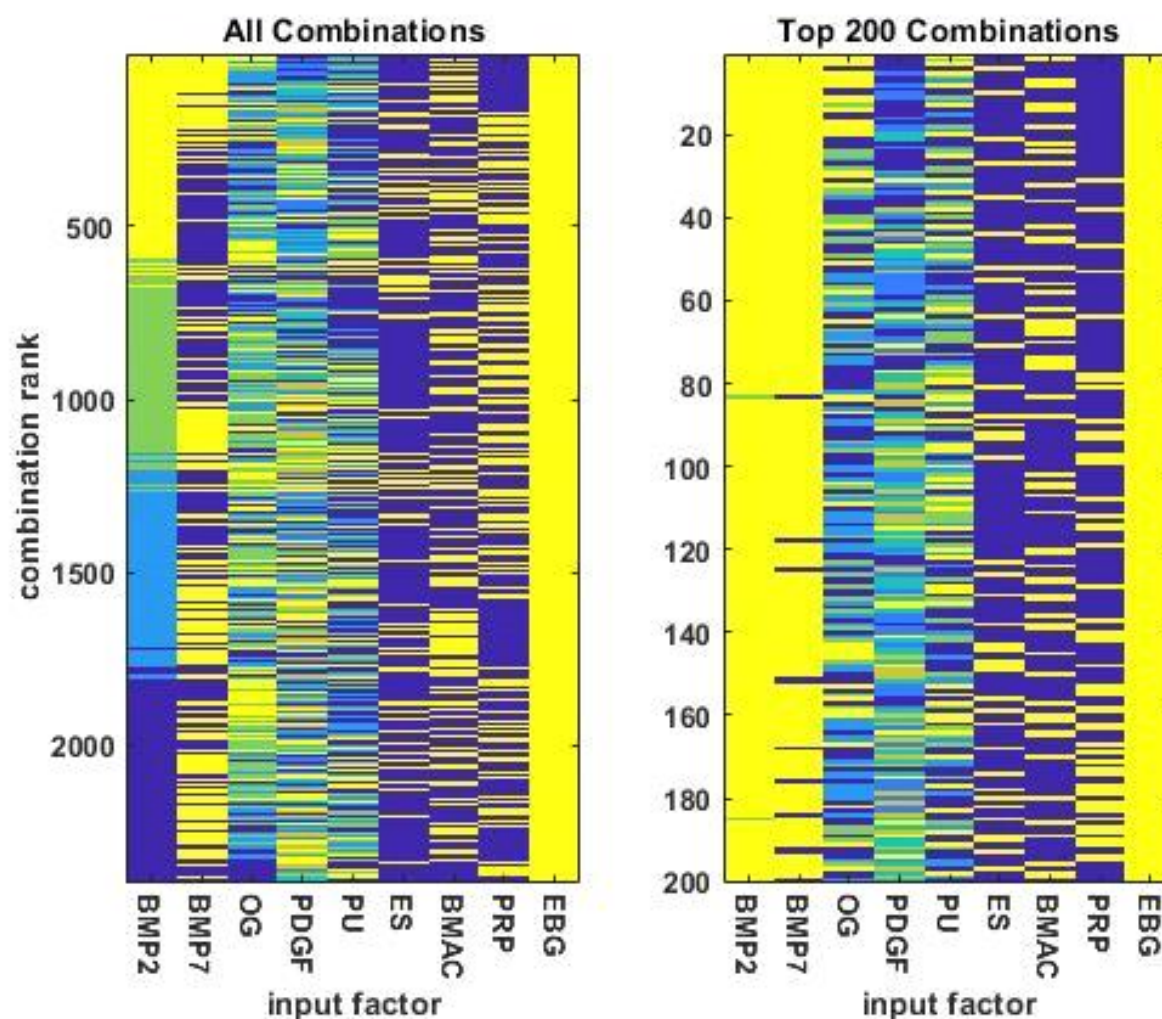


Figure 4. The combinations of orthobiologic input factors ordered by predicted efficacy. The combinations are presented in descending order, so the combinations are most / least effective at the top / bottom of either plot. The plot on the left shows all 2400 combinations while the plot on the right shows the top 200 combinations. The factors ES, BMAC, and PRP are either present or absent; EBG is always present. The factor BMP7 takes 2 levels; BMP2, OG, and PU each take 4 levels; and PDGF takes 5 levels. Input factor levels varied over a broad range but were normalized into [0, 1] for purposes of illustration. In the images, yellow and deep blue correspond to 1 and 0, respectively.

REFERENCES

1. Wang, M., N. Yang, and X. Wang, *A review of computational models of bone fracture healing*. Med Biol Eng Comput, 2017. **55**(11): p. 1895-1914.
2. Smith, B., T. Goldstein, and C. Ekstein, *Biologic adjuvants and bone: current use in orthopedic surgery*. Curr Rev Musculoskelet Med, 2015. **8**(2): p. 193-9.
3. Goodfellow, I., Bengio, Y, Courville, A, *Deep Learning*. 2016: MIT press.
4. LeCun, Y., Y. Bengio, and G. Hinton, *Deep learning*. Nature, 2015. **521**(7553): p. 436-44.
5. Schmidhuber, J., *Deep learning in neural networks: An overview*, in *Neural Networks*. 2015. p. 61-85.
6. Cao, C., et al., *Deep Learning and Its Applications in Biomedicine*. Genomics Proteomics Bioinformatics, 2018. **16**(1): p. 17-32.
7. Mamoshina, P., et al., *Applications of Deep Learning in Biomedicine*. Mol Pharm, 2016. **13**(5): p. 1445-54.
8. Wainberg, M., et al., *Deep learning in biomedicine*. Nat Biotechnol, 2018. **36**(9): p. 829-838.
9. Chan, S. and E.L. Siegel, *Will machine learning end the viability of radiology as a thriving medical specialty?* Br J Radiol, 2019. **92**(1094): p. 20180416.
10. Choy, G., et al., *Current Applications and Future Impact of Machine Learning in Radiology*. Radiology, 2018. **288**(2): p. 318-328.
11. Dreyer, K.J. and J.R. Geis, *When Machines Think: Radiology's Next Frontier*. Radiology, 2017. **285**(3): p. 713-718.
12. Munk, P.L., *A Brave New World Without Diagnostic Radiologists? Really?* Can Assoc Radiol J, 2018. **69**(2): p. 119.
13. Cho, S., Won, HH, *Cancer classification using ensemble of neural networks with multiple significant gene subsets*. Applied Intelligence, 2007. **26**(243).
14. Khan, J., et al., *Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks*. Nat Med, 2001. **7**(6): p. 673-9.
15. Guler, I., Ubeyli, ED, *ECG beat classifier designed by combined neural network model*. Pattern Recognition, 2004. **38**(199).
16. Mantzaris, D., Anastassopoulos, G, Lymberopoulos, D, *Medical disease prediction using Artificial Neural Networks*. 8th IEEE International Conference on Bioinformatics and BioEngineering, 2008. **10**.
17. Lekha, S. and S. M, *Real-Time Non-Invasive Detection and Classification of Diabetes Using Modified Convolution Neural Network*. IEEE J Biomed Health Inform, 2018. **22**(5): p. 1630-1636.
18. Zecchin, C., et al., *Jump neural network for real-time prediction of glucose concentration*. Methods Mol Biol, 2015. **1260**: p. 245-59.
19. Choi, E., et al., *Doctor AI: Predicting Clinical Events via Recurrent Neural Networks*. JMLR Workshop Conf Proc, 2016. **56**: p. 301-318.
20. Nguyen, P., et al., *DeepPr: A Convolutional Net for Medical Records*. IEEE J Biomed Health Inform, 2017. **21**(1): p. 22-30.
21. Weng, S.F., et al., *Can machine-learning improve cardiovascular risk prediction using routine clinical data?* PLoS One, 2017. **12**(4): p. e0174944.
22. Grigsby, J., R. Kookan, and J. Hershberger, *Simulated neural networks to predict outcomes, costs, and length of stay among orthopedic rehabilitation patients*. Arch Phys Med Rehabil, 1994. **75**(10): p. 1077-81.
23. Lin, C.C., et al., *Comparison of artificial neural network and logistic regression models for predicting mortality in elderly patients with hip fracture*. Injury, 2010. **41**(8): p. 869-73.
24. Griffin, X.L., et al., *Electromagnetic field stimulation for treating delayed union or non-union of long bone fractures in adults*. Cochrane Database Syst Rev, 2011(4): p. CD008471.

25. Rutten, S., et al., *Enhancement of Bone-Healing by Low-Intensity Pulsed Ultrasound: A Systematic Review*. JBJS Rev, 2016. **4**(3).
26. Perrone, M., Cooper, LN, *When networks disagree: Ensemble methods for hybrid neural networks* in *How We Learn; How We Remember: Toward An Understanding Of Brain And Neural Systems: Selected Papers of Leon N Cooper* 1995, World Scientific. p. 342-358.
27. Wang, Z., C.C. Clark, and C.T. Brighton, *Up-regulation of bone morphogenetic proteins in cultured murine bone cells with use of specific electric fields*. J Bone Joint Surg Am, 2006. **88**(5): p. 1053-65.