Machine learning-facilitated optimization of nano-drug formulations

Sihan Dong and Emmanuel A. Ho
Laboratory for Drug Delivery and Biomaterials, School of Pharmacy, University of Waterloo

INTRODUCTION

Developing nanoparticle (NP) formulations through trial and error is time-consuming and expensive. As a proof-of-concept, we will use machine learning (ML) with Python to determine the mathematical relationships between four physicochemical properties (lactic acid to glycolic acid ratio (LA/GA), polymer molecular weight (MW), polymer to drug ratio, and drug lipophilicity (logP)) and the corresponding drug encapsulation efficiency (EE%) and therapeutic efficacy with a goal of optimizing the NP formulation compositions.

OBJECTIVES

- Create 48 NPs that are varied in poly (lactic-co-glycolic) acid (PLGA) LA/GA ratio, PLGA MW, polymer to drug ratio, and drug logP.
- Collect corresponding EE% and therapeutic efficacy (i.e., half-maximal inhibitory concentration (IC50)) for each NP.
- Feed collected data into partial least square regression (PLSR), artificial neural network (ANN), and reinforcement learning (RL) models.
- Compare predicted data and observed (i.e., true) data for evaluation of prediction accuracy.

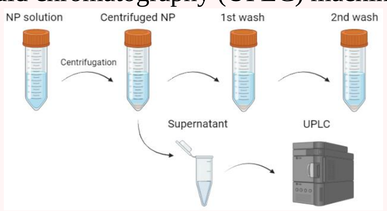
MATERIALS

- Drugs: Docetaxel (DTX), carboplatin (CPT), doxorubicin (DOX)
- Polymer: PLGA, polyvinyl alcohol (PVA)
- Cell line: human epithelial ovarian cancer cells (OVCAR-3)

METHODS

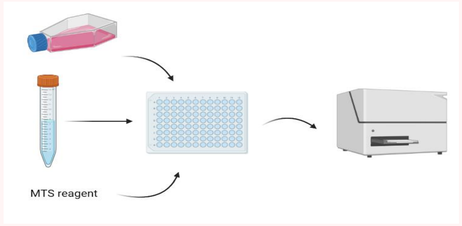
***In-vitro**:*

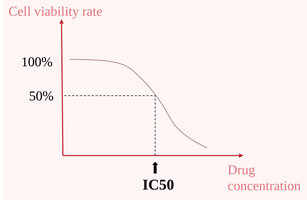
PLGA NPs are made by double-emulsion method and are further characterized through size and zeta potential measurements using the Zetasizer. The EE% is measured indirectly using ultra performance liquid chromatography (UPLC) machine.



$$EE\% = \frac{Total\ \text{amount of}\ drugs - drugs\ in\ \text{the}\ supernatant}{Total\ \text{amount of}\ drugs} \times 100\%$$

MTS cytotoxicity assays are conducted to calculate IC50 values from cell viability curves.





$$\text{Cell viability\%} = \frac{Treated - blank}{Negative\ control - blank} \times 100\%$$

In-silico:

- PLSR

EE% or Therapeutic efficacy (i.e., IC50) = e + β_1 * (LA/GA ratio) + β_2 * (polymer to drug ratio) + β_3 * MW + β_4 * logP

- ANN

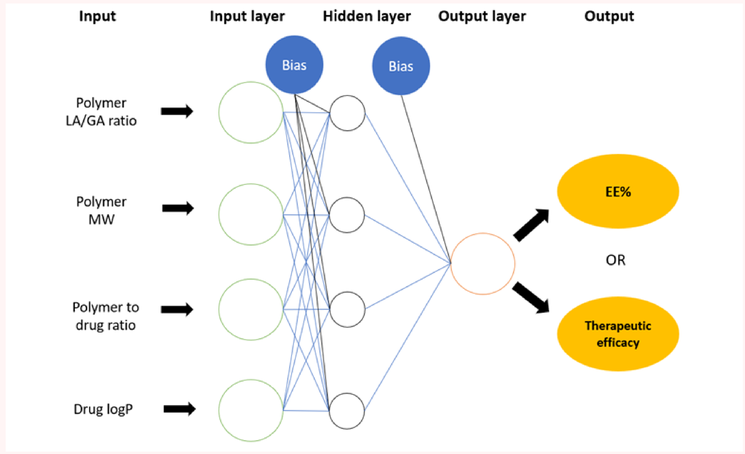


Figure 1. ANN structure schematic adapted from Shalaby et al. (2014)[1]

- RL

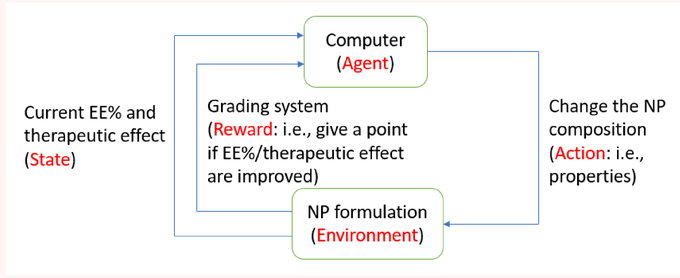


Figure 2. RL circuit schematic adapted from Galatzer-Levy et al. (2018)[2]

RESULTS

NP	Average size (nm)	Average PDI	Zeta potential (mV)	EE%
1mg DTX in 20mg 50:50 5000Da PLGA	191.9±2.194	0.097±0.030	-22.9±0.45	49.27±0.20
1mg CPT in 20mg 50:50 5000Da PLGA	194.1±4.236	0.080±0.007	-24.7±1.50	2.62±2.09
3mg CPT in 25mg 50:50 2000Da PLGA	189.6±10.49	0.163±0.048	-30.1±2.65	26.69±2.17
3mg CPT in 25mg 50:50 5000Da PLGA	213.8±8.316	0.132±0.017	-30.9±2.78	12.60±1.39

Table 1. NP size, PDI, zeta potential, and EE% for different NP formulations.

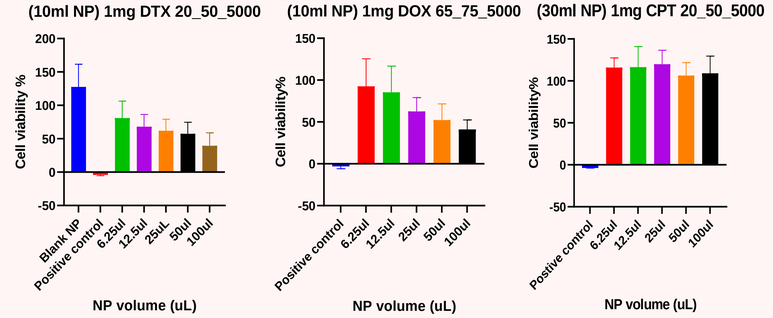


Figure 3. A). Cell viability curve for 1mg DTX in 20mg 50:50, 5000Da PLGA. B). Cell viability curve for 1mg DOX in 65mg 75:25, 5000Da PLGA. C). Cell viability curve for 1mg CPT in 20mg 50:50, 5000Da PLGA. Positive controls: Triton X-100.

CONCLUSION

EE% and therapeutic efficacy collected from *in-vitro* experiments will be inputted and memorized by the computer program. Through repeated data training, the programs will be able to predict the encapsulation efficiency and therapeutic efficacy from provided NP composition details.

REFERENCES

1. Shalaby, K. et al. Determination of factors controlling the particle size and entrapment efficiency of noscapine in PEG/PLA nanoparticles using artificial neural networks. IJN 4953 (2014) doi:10.2147/IJN.S68737
2. Galatzer-Levy, I. R., Ruggles, K. V. & Chen, Z. Data Science in the Research Domain Criteria Era: Relevance of Machine Learning to the Study of Stress Pathology, Recovery, and Resilience. *Chronic Stress* **2**, 247054701774755 (2018).