UNIVERSITI TEKNOLOGI MARA

PREDICTION OF NOVEL DOPING AGENT USING AN *IN SILICO* MODEL THAT INTEGRATES CHEMICAL, BIOLOGICAL AND PHENOTYPIC DATA

NURUL AMALINA BINTI JAMIL

Dissertation submitted in partial fulfilment of the requirements for the Bachelor of Pharmacy (Hons.)

Faculty of Pharmacy

JULY 2016

ACKNOWLEDGEMENTS

Assalamualaikum w.b.t.

First of all, Alhamdullilah, I would like to thank Allah S.W.T for the given opportunities for me to learn something new and for the courage to start this project. Secondly, I would like to thank my supervisor, Dr. Fazlin Mohd Fauzi for her guidance and patience, repeating the important information for this project and her constant commitment in correcting the thesis. I would also like to thank my parents for their constant motivation and encouragement, always believing that I can do more than what I think I am capable of. I would also like to thank my project members, Ain and Artif for reminding me what I have forgotten, for correcting me if I misunderstood something and being together through any hardships. To all my friends, thank you for lending me your ears and shoulders when I needed them. The constant supports from all these people will never be forgotten.

TABLE OF CONTENTS

ABSTRACT

The anticipation of novel doping agents is usually not quick enough to prevent their misuse among athletes. Current techniques e.g. High performance Liquid Chromatography (HPLC) to detect novel doping agents usually involves a time-consuming process. Computational method or *in silico* model offers a quick and accurate method to detect potential doping agents. The *in silico* model used in this study integrates chemical, biological and phenotypic data to identify novel doping agent. In this study, two different training sets, termed as biological and phenotypic were compiled and three molecular descriptors (MACCS, ECFP4, FCFP4) and two machine learning algorithms (Naïve Bayes and Decision Tree) were employed to build the predictive models. These models were validated using five-fold cross validation to evaluate their predictive power. Both biological and phenotypic models were then combined using Joint Belief model. Sensitivity or the ability of the model to identify doping agent accurately between individual and combination models were compared. This study found that the combination model predicts better compared to the individual models with highest sensitivity of 0.962. Four compounds, Meldonium, Mitragynine, Pipradol and Synephrine were then tested in both individual and combination models. All four compounds were predicted as non-doping agents using biological models and were predicted as doping agents using phenotypic models. Combination of both biological and phenotypic models predicted all four compounds to be doping agents. The statement are supported by study done on Meldonium which is an anti-ischemic drug that was recently banned due to its effect on myocardial function that improves athlete's endurance. Mitragynine which is listed in WADA Monitoring Program, has also the potential as novel doping agent due to its binding activity on κ-opioid receptors. The activation of κ-opioid receptors produce analgesic effect which also aids in endurance of athletes. Hence, the use of *in silico* predictive model allowed the detection of novel doping agents quickly and can prevent the spreading of them among athletes.

CHAPTER 1

INTRODUCTION

1.1 History of doping

On October 2014, the nation was shocked with the news that our top ranked badminton player, Dato' Lee Chong Wei failed the drug test at the World Championships in Copenhagen, Denmark. As a consequence, he was given an eight-months bans from any competition (Paul, n.d.). It was reported that he took food supplement that contained doping agent. The doping agent was revealed to be Dexamethasone, a substance that is listed on the World Anti-Doping Agency (WADA) International Standard of Prohibited Substances and Methods list. Dexamethasone is an anti-inflammatory steroid agent that is not illegal outside of competition but is ban during the competition due to its performance-enhancing effect. In a clinical setting, Dexamethasone is a common steroid drug prescribed for conditions such as skin allergies and rheumatic problems. However due to their effect in increasing an athlete's performance during the competition, it was banned.

According to ("World Anti-Doping Agency," n.d.)., doping came from the Dutch word *dop*, an alcoholic beverage that was used by their ancient warriors in order to increase their ability and skills during battle. The history of doping dated back hundreds of years ago, during ancient Greek games where the athletes drank wine and brandy