6-9-2017

Data-Driven Design of an Ebola Therapeutic

Robert A. Lodder
University of Kentucky, r.lodder@uky.edu

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 108C (2017) 1612–1621

Procedia Computer Science

International Conference on Computational Science, ICCS 2017, 12-14 June 2017, Zurich, Switzerland

Data-Driven Design of an Ebola Therapeutic

Robert A. Lodder[1]

[1]University of Kentucky
Lexington, KY
Lodder@uky.edu

Abstract

Formulation is very important in drug delivery. The wrong formulation can render a drug product useless. The amount of preclinical (animal and in vitro) work that must be done before a new drug candidate can be tested in humans can be a problem. The cost of these cGxP studies is typically $3-$5 million. If the wrong drug product formulation is tested, new iterations of the formulation must be tested with additional costs.

Data-driven computational science can help reduce this cost. In the absence of existing human exposure, a battery of preclinical tests must be performed in at least two species before FDA will permit testing in humans. However, for many drugs (such as those beginning with natural products) there is a history of human exposure. In these cases, computer modeling of a population to determine human exposure may be adequate to permit phase 1 studies with a candidate formulation in humans.

The CDC's National Health and Nutrition Examination Survey (NHANES) is a program of studies designed to assess the health and nutritional status of adults and children in the United States. The survey is unique in that it combines interviews and physical examinations including laboratory results. The NHANES database can be mined to determine exposure to a food additive, and early human formulation testing conducted at levels beneath those to which the US population is ordinarily exposed through food. These data can be combined with data mined from international chemical shipments to validate an exposure model. This paper describes the data driven formulation testing process using a new candidate Ebola treatment that, unlike vaccines, can be used after a person has contracted the disease. This drug candidate's mechanism of action permits it to be potentially used against all strains of the virus, a characteristic that vaccines might not share.

Peer-review under responsibility of the scientific committee of the International Conference on Computational Science

Keywords: formulation, toxicology, ICH, FDA, pharmaceutical development.

1 Introduction

Data-driven computational science has found many applications in drug design. Molecular data are commonly used to design new drug molecules. Engineering process simulations guide the development of the Chemistry, Manufacturing, and Controls (CMC) section of Investigational New

1877-0509
Peer-review under responsibility of the scientific committee of the International Conference on Computational Science
10.1016/j.procs.2017.05.127

Drug (IND) applications filed at FDA. Computer simulations can also guide the design of human clinical trials.

Recently, a study of molecular modeling of structural and conformational relationships in tRNA complexation with chloroethyl nitrosourea derivatives was published (Agarwal, 2017). These derivatives are used as chemotherapeutic agents against cancers including brain tumors, Hodgkin's disease, small cell lung cancer, and malignant melanoma. Molecular modeling has been recently used to analyze hydrogen bonding to Thr120 and Thr124 for thiadiazolodiazepine analogs used as neuromuscular blocking agents (El-Subbagh, 2017). Poureshghi et al. have used spectroscopic techniques combined with molecular modeling methods to study the interaction between lamotrigine and human serum albumin (Poureshghi, 2017) Phenylpicolinamide derivatives have been modeled using SURFLEX-DOCK module of the SYBYL package (Zhu, 2016)

Boehling et al. have published a simulation experiment using the discrete element method to study the influence of process parameters including spray rate, number of nozzles, drum rotation rate, and drum load on the inter-tablet coating variation (Boehling, Analysis of large-scale tablet coating: Modeling, simulation and experiments, 2016). Boehling et al. have also studied the scale up process for spray coating computationally. Industrial scale up can be a complicated process and sometimes fails (Boehling, 2016b). Sensitivity analyses of a simulated pharmaceutical direct compaction process using Sobol indices and based on steady-state gains and the frequency response of a planned production plant were conducted by Rehrl et al. (Rehrl, 2017). The continuous manufacturing of artemisinin has been studied by process simulations (Jolliffe, 2016).

Of the three major applications of simulation in pharmaceutical development, clinical trial simulations are perhaps the most challenging. The challenge arises from the need to model the behavior of a small group of individual human subjects. A company working to get a single drug to market can expect to spend at least $350 million before the medicine is approved for sale. Because so many new drug candidates fail, large pharmaceutical companies that are working on dozens of drug projects simultaneously spend $5 billion per new drug (Herper, 2013).

Clinical trial simulations can reduce the risk of conducting clinical trials and thus reduce the costs. Mileham et al. have published on risk modeling for clinical trial patient accrual (Mileham, 2016). Banks et al. have conducted simulations of clinical trials with missing data (Banks, 2017). Basu used modeling to resolve differences between two recent randomized clinical trials intended to measure the benefits and harms of blood pressure treatments in patients with cardiovascular disease (Basu, 2017). Finally, Bayesian approaches are being increasingly applied to design and analysis of clinical trials. Wang et al. took a Bayesian approach to an interim futility analysis of a clinical trial (Wang, 2016).

2 Purpose

The purpose of this study is to use simulations and modeling in a new way: to obviate the need for extensive preclinical formulation and toxicology studies, and to thereby speed a candidate Ebola therapeutic to the clinic.

To achieve this purpose, data from the Centers for Disease Control (CDC) National Health and Nutrition Examination Survey (NHANES) are combined with data on shipments from food ingredient manufacturers to estimate exposure of the US population through foods to a candidate pharmaceutical formulation ingredient, beta cyclodextrin (BCD). Formulation is important in drug delivery. A poor formulation can render a drug product useless. A large amount of preclinical research must be performed before a new drug candidate can be tested in humans. The cost of these cGxP studies is typically $3-$5 million. If the wrong drug product formulation is tested preclinically, new iterations of the formulation must be retested with additional costs. BSN389 Is a diterpene antibiotic designed for treatment of Ebola virus infections. The drug arrests viral replication in the body while antibody

therapies are administered to eliminate the virus. However, the drug is poorly soluble in water, reducing circulating blood levels of the drug. The goal of formulation is to increase the solubility and bioavailability.

Data-driven computational science can help reduce the cost. In the absence of existing human exposure, a battery of tests involving acute and chronic toxicology, cardiovascular, central nervous system, and respiratory safety pharmacology must be performed in at least two species before FDA will permit testing in humans. However, for many compounds (such as those beginning with natural products) there is a history of human exposure. In these cases, computer modeling of a population to determine human exposure may be adequate to permit phase 0-1 studies with a candidate formulation in humans.

The CDC's National Health and Nutrition Examination Survey (NHANES) is a set of studies designed to evaluate the health and nutritional status of adults and children in the United States. NHANES is unique in that it combines interviews and dietary information with physical examinations including laboratory results. The NHANES database can be mined to determine the distribution of exposures to a food additive, and early human formulation testing conducted at levels beneath those to which the US population is ordinarily exposed through food. These data can be combined with data mined from international chemical shipments to validate an exposure model. This paper describes the data driven formulation testing process using a new candidate Ebola treatment that, unlike vaccines, can be used after a person has contracted the disease. This BSN389's mechanism of action permits it to be potentially used against all strains of the virus, a characteristic that vaccines might not share.

3 β-Cyclodextrin Uses in Food and Pharmaceuticals

Cyclodextrins (CDs) are available in three common varieties: α-cyclodextrin, β-cyclodextrin and γ-cyclodextrin, which are together called the first generation (or parent) cyclodextrins. These cyclodextrins are composed of six (α), seven (β) and eight (γ) -(1,4)-linked glycosyl units formed into a ring. This ring-shaped molecule is hydrophilic on the outside (so the CD can dissolve in water) and has a nonpolar cavity inside, which provides a hydrophobic environment. Because of this hydrophobic cavity, cyclodextrins are able to form inclusion complexes with a variety of hydrophobic guest molecules. The fact that β cyclodextrin is used in both food and pharmaceuticals complicates BCD exposure estimates, and means that manufacturers must be careful when adding what seems to be a small amount of BCD to a product because that small amount may be enough to drive a consumer already consuming BCDs from other sources above the Acceptable Daily Intake limit.

CDs are well understood from a regulatory point of view, and a monograph for BCD appeared decades ago in both the US Pharmacopoeia/National Formulary and the European Pharmacopoeia (Del Valle, 2004).

3.1 β-Cyclodextrin

The oral LD50 in the rat is greater than 5000 mg/kg, and the i.v. LD50 in the rat is between 450 and 790 mg/kg. BCD is less irritating than α-cyclodextrin after i.m. injection. BCD binds cholesterol and only very small amounts (1–2%) are absorbed in the upper intestinal tract after oral administration.

3.2 Applications of β-Cyclodextrin

In a CD each guest molecule is effectively microencapsulated. This encapsulation is employed to produce favorable changes in the chemical and physical properties of the guest molecules, including.

• Improving the solubility of guest molecule.

• Immobilizing very volatile substances.
• Turning liquid forms into powders.
• Modifying the chemical reactivity of guest molecules.
• Stabilizing photosensitive or oxygen-sensitive molecules.
• Protecting molecules from degradation by microorganisms.
• Masking of bad color, smell or taste.

These types of uses of cyclodextrins and CD derivatives motivate their many applications in food and pharmaceuticals, and increase the possibility of human exposure.

3.3 Foods and flavors

CDs are used in food preparations for flavor protection or flavor delivery. In foods CDs form inclusion complexes with many different types of molecules including lipids, flavors and colors. Most natural and artificial flavors are volatile oils or liquids, and complexing them with CDs provides an alternative to other encapsulation technologies used for flavor protection. CDs are also employed as process aids, e.g., to remove cholesterol from products like milk, butter and eggs. CDs can have a texture-improving effect on pastry and on meat products. Other applications include reducing bitterness, unpleasant smells and tastes, and stabilizing flavors subjected to long-term storage. Emulsions like mayonnaise, margarine or butter creams can be stabilized with α-cyclodextrin. As a result, CDs are found in many foods to which consumers are exposed.

Cyclodextrins are useful in pharmaceuticals for many of the same reasons that they are useful in foods. For example, the addition of α- or β-cyclodextrin increases the water solubility of poorly water-soluble drugs. In some cases improving solubility improves bioavailability, increasing the effectiveness of the drug and enabling a reduction in dose. Cyclodextrins can be used analytically or preparatively in the separation of drug stereoisomers by the formation of β-cyclodextrin inclusion complexes (Armstrong, 1986). As a result, CDs are found in pharmaceuticals to which consumers are exposed.

3.4 Assessment of β-Cyclodextrin Use

An assessment of the consumption of BCD by the U.S. population resulting from the approved uses of BCD was conducted. Estimates for the intake of BCD were based on the approved food uses and maximum use level in conjunction with food consumption data included in the National Center for Health Statistics' (NCHS) 2009-2010, 2011-2012, and 2013-2014 National Health and Nutrition Examination Surveys (NHANES) (Prevention, 2005) (USDA, 2012); (Bodner-Montville J, 2006). Calculations for the mean and 90th percentile intakes were performed for representative approved food uses of BCD combined (see Appendix for lists of food codes used). The intakes were reported for the following population groups:

- infants, age 0 to 1 year
- toddlers, age 1 to 2 years,
- children, ages 2 to 5 years,
- children, ages 6 to 12 years,
- teenagers, ages 13 to 19 years,
- adults, ages 20 years and up,
- total population (all age groups combined, excluding ages 0-2 years)

4 Food Consumption Survey Data

4.1 Survey Description

The most recent National Health and Nutrition Examination Surveys (NHANES) for the years 2013-2014 are available for public use. NHANES are conducted as a continuous, annual survey, and are released in 2-year cycles. In each cycle, approximately 10,000 people across the U.S. completed the health examination component of the survey. Any combination of consecutive years of data collection is a nationally representative sample of the U.S. population. It is well established that the length of a dietary survey affects the estimated consumption of individual users and that short-term surveys, such as the typical 1-day dietary survey, overestimate consumption over longer time periods. Because two 24-hour dietary recalls administered on 2 non-consecutive days (Day 1 and Day 2) are available from the NHANES 2009-2010, 2011-2012, and 2013-2014 surveys, these data were used to generate estimates for the current intake analysis.

The NHANES provide the most appropriate data for evaluating food-use and food-consumption patterns in the United States, containing 2 years of data on individuals selected via stratified multistage probability sample of civilian non-institutionalized population of the U.S. NHANES survey data were collected from individuals and households via 24-hour dietary recalls administered on 2 non-consecutive days (Day 1 and Day 2) throughout all 4 seasons of the year. Day 1 data were collected in-person in the Mobile Examination Center (MEC), and Day 2 data were collected by telephone in the following 3 to 10 days, on different days of the week, to achieve the desired degree of statistical independence. The data were collected by first selecting Primary Sampling Units (PSUs), which were counties throughout the U.S. Small counties were combined to attain a minimum population size. These PSUs were segmented and households were chosen within each segment. One or more participants within a household were interviewed. Fifteen PSUs are visited each year. For example, in the 2009-2010 NHANES, there were 13,272 persons selected; of these 10,253 were considered respondents to the MEC examination and data collection. 9754 of the MEC respondents provided complete dietary intakes for Day 1 and of those providing the Day 1 data, 8,405 provided complete dietary intakes for Day 2.

Sample weights were incorporated with NHANES surveys to compensate for the potential under-representation of intakes from specific population groups as a result of sample variability due to survey design, differential non-response rates, or other factors, such as deficiencies in the sampling frame (Prevention, 2005) (USDA, 2012).

4.2 Statistical Methods

Consumption data from individual dietary records, detailing food items ingested by each survey participant, were collated by computer in Matlab and used to generate estimates for the intake of BCD by the U.S. population. Estimates for the daily intake of BCD represent projected 2-day averages for each individual from Day 1 and Day 2 of NHANES data; these average amounts comprised the distribution from which mean and percentile intake estimates were produced. Mean and percentile estimates were generated incorporating sample weights in order to provide representative intakes for the entire U.S. population. “All-user” intake refers to the estimated intake of BCD by those individuals consuming food products containing BCD. Individuals were considered users if they consumed 1 or more food products containing BCD on either Day 1 or Day 2 of the survey.

5 Food Usage

5.1 Food Data

Food Type	Max. Use
baked goods prepared from dry mixes	2.0%
breakfast cereal	2.0%
chewing gum	2.0%
gelatins and puddings	1.0%
dry mix for soups	0.2%
flavored coffee and tea	1.0%
compressed candies	2.0%
processed cheese products	1.0%
flavored savory snacks and crackers	0.5%
dry mix for beverages	1.0%

Table 1. Maximum Permitted BCD Use in Foods (by wt.)

Food codes representative of each approved use (Table 1) were chosen from the Food and Nutrition Database for Dietary Studies (FNDDS) for the corresponding biennial NHANES survey.

5.2 Food Survey Results

The estimated "all-user" total intakes of BCD from all approved food uses of BCD in the U.S. by population group is summarized in Table 2.

Population Group	N users	N population	% Users	Mean mass (kg)	Mean EDI (g)	90th % EDI (g)	Mean EDI (g/kg)	90th % EDI (g/kg)
ages 0-1	215	379	56.73	7.97	0.0018	0.0043	0.0002	0.0005
ages 1-2	184	290	63.45	13.69	0.0077	0.0166	0.0006	0.0012
ages 2-5	718	1196	60.03	15.91	0.0080	0.0165	0.0005	0.0010
ages 6-12	929	1564	59.40	37.16	0.0121	0.0262	0.0003	0.0007
ages 13-19	727	1239	58.68	68.80	0.0159	0.0301	0.0002	0.0004
ages 20 and up	2721	5769	47.17	81.46	0.0197	0.0589	0.0002	0.0007
ages 2 and up	5095	9768	52.16	66.45	0.0160	0.0363	0.0002	0.0005

Table 2. Estimated "All-user" Daily Intake (EDI) of BCD in Targeted Foods by Population Group (2013-2014 NHANES Data)

Of course, every food category in which BCD is approved for use does not necessarily incorporate BCD into every product at the maximum approved use level. As a result, the values in Table 2 are corrected using the total amount of BCD consumed in food in the United States during the period of the survey. The correction was derived from US population numbers and market research on the global β cyclodextrin industry (Research, 2016). The US population numbers by year for 2010, 2012, and 2014 are 308.11 million , 312.86 million, and 317.68 million, respectively (US population numbers, 2017).

To derive the correction factor the US population number for 2014 was multiplied by the fraction of people aged two and up in the US consuming the targeted food codes:

317680000 persons x 0.5216 = 1.65702e+008, or 165.702 million BCD consumers

971 tons of BCD were consumed in US foods in 2014 (see Figure 1). This number was converted to grams and divided by the number of BCD consumers in the United States to give:
971000000/165702000 = 5.859 g/consumer/yr

Dividing the grams per consumer per year by 365 gives grams per consumer per day, or 0.01605. To derive the correction factor the maximum g/consumer/day for ages 2+ in the NHANES table is divided by the actual g/day, 0.5617/0.01605 = 35.0. The estimated maximum exposures provided by the 2013-2014 NHANES were divided by 35 to get estimated actual exposures in Table 2.

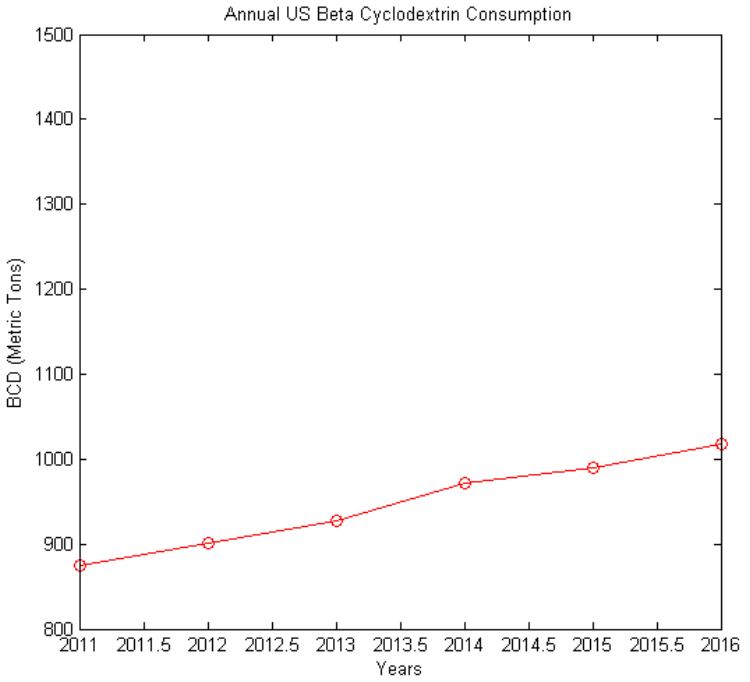

Figure 1. Annual US β cyclodextrin consumption by year (metric tons)

BCD exposure from foods is forecast to increase in the near future. There is a correlation between BCD total consumption and exposure estimated from NHANES. Linear modeling of the NHANES data predicts exposure will reach 19 mg by 2020 in the age 2 and up group (see Figure 2).

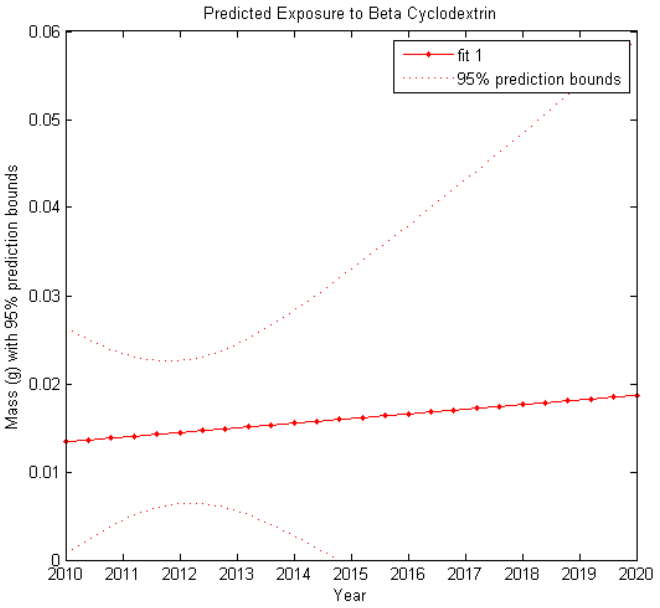

Figure 2. The mass of BCD consumed by users in the US is forecast to increase to 19 mg/person/day in the age 2 and up group by the year 2020.

Much of the increased exposure to BCD seems to come from an increasing number of food codes in categories permitted to incorporate BCD. Figure 3 shows the number of food codes increasing through 2020, which is also correlated to the amount of BCD consumed.

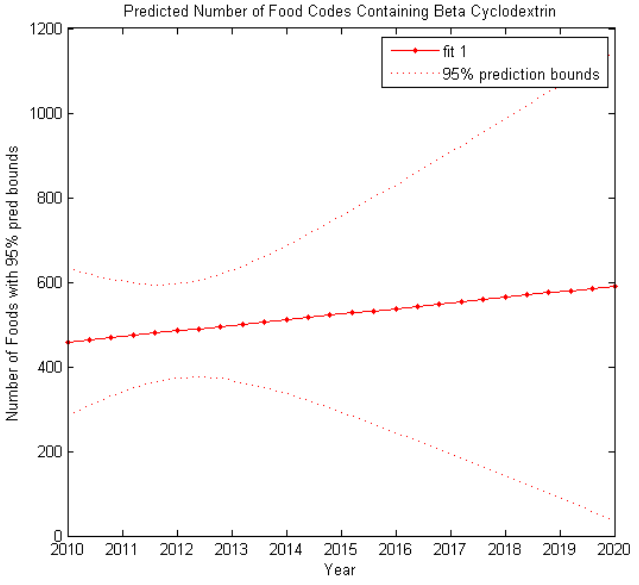

Figure 3. Every year more food codes are added to categories of foods permitted to incorporate BCD. By 2020 nearly 600 food codes containing BCD may be listed in the FNDDS.

6 Conclusions

In summary, 52.2% of the total U.S. population of 2+ years was identified as consumers of BCD from the approved food uses. The mean intakes of BCD by the all BCD consumers ("all-user") from all approved food uses were estimated to be 16 mg/person/day or 0.2 mg/kg body weight/day. The heavy consumer (90th percentile all-user) intakes of BCD from all approved food-uses were estimated to be 36.3 mg/person/day or 0.5 mg/kg body weight/day. The initial human clinical studies of BSN389 will use 1.5 µg of BCD. This is four orders of magnitude less than the expected daily intake from food uses, and far less than the amount required to take the average consumer from the 50th to the 90th percentile. For this reason, use of BCD in the BSN 389 formulation is an insignificant addition to daily intake and should be safe for subjects in the trial.

7 Support

The project described was supported by the NIH National Center for Advancing Translational Sciences through grant number UL1TR001998. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575.

References

Agarwal, Shweta, et al. "Structural-conformational aspects of tRNA complexation with chloroethyl nitrosourea derivatives: A molecular modeling and spectroscopic investigation." Journal of Photochemistry and Photobiology B: Biology 166 (2017): 1-11.

Armstrong, Daniel W., et al. "Separation of drug stereoisomers by the formation of beta-cyclodextrin inclusion complexes." Science 232 (1986): 1132-1136.

Banks, Harvey Thomas, Shuhua Hu, and Eric Rosenberg. "A dynamical modeling approach for analysis of longitudinal clinical trials in the presence of missing endpoints." Applied Mathematics Letters 63 (2017): 109-117.

Basu, Sanjay, Jeremy B. Sussman, and Rod A. Hayward. "Detecting Heterogeneous Treatment Effects to Guide Personalized Blood Pressure Treatment: A Modeling Study of Randomized Clinical Trials." Annals of Internal Medicine Jan. 3. DOI: 10.7326/MlG-1756 (2017).

Bodner-Montville J, J. K.C. Ahuja, L. A. Ingwersen, E. S. Haggerty, C. W. Enns and B. P. Perloff 2006. USDA Food and Nutrient Database for Dietary Studies: Released on the web. Journal of Food Composition and Analysis, 19(Supplement 1): S100-S107.

Boehling, P., et al. "Analysis of large-scale tablet coating: Modeling, simulation and experiments." European Journal of Pharmaceutical Sciences 90 (2016a): 14-24.

Boehling, P., et al. "Simulation of a tablet coating process at different scales using DEM." European Journal of Pharmaceutical Sciences 93 (2016b): 74-83.

CDC 2006. Analytical and Reporting Guidelines: The National Health and Nutrition Examination Survey (NHANES). National Center for Health Statistics, Centers for Disease Control and Prevention; Hyattsville, Maryland. Available from: http://www.cdc.gov/nchs/data/nhanes/nhanes_03_04/nhanes_analytic_guidelines_dec_2005.pdf

Del Valle, EM Martin. "Cyclodextrins and their uses: a review." Process biochemistry 39.9 (2004): 1033-1046.

El-Subbagh, Hussein I., et al. "Thiadiazolodiazepine analogues as a new class of neuromuscular blocking agents: Synthesis, biological evaluation and molecular modeling study." European Journal of Medicinal Chemistry 126 (2017): 15-23.

Global and Chinese Beta Cyclodextrin Industry, 2016 Market Research Report, Prof Research Nov. 2016.

Herper, Matthew. The Cost Of Creating A New Drug Now $5 Billion, Pushing Big Pharma To Change. Forbes. Retrieved Sep 12 2013.

http://www.forbes.com/sites/matthewherper/2013/08/11/how-the-staggeringcost-of-inventing-new-drugs-is-shaping-the-future-of-medicine/

Jolliffe, Hikaru G., and Dimitrios I. Gerogiorgis. "Process modelling and simulation for continuous pharmaceutical manufacturing of artemisinin." Chemical Engineering Research and Design 112 (2016): 310-325.

Mileham, Kathryn F., and Edward S. Kim. "Who's Eligible Anyway? Risk Modeling for Clinical Trial Accrual." Clinical Cancer Research 22.22 (2016): 5397-5399.

Poureshghi, Fatemeh, et al. "Interaction of an antiepileptic drug, lamotrigine with human serum albumin (HSA): Application of spectroscopic techniques and molecular modeling methods." Journal of Photochemistry and Photobiology B: Biology 166 (2017): 187-192.

Rehrl, Jakob, et al. "Sensitivity analysis of a pharmaceutical tablet production process from the control engineering perspective." International Journal of Pharmaceutics 517.1 (2017): 373-382.

US population numbers, http://www.multpl.com/united-states-population/table. Retrieved Jan. 30, 2017

USDA 2012. What We Eat In America (WWEIA), NHANES: overview. Available from: http://www.ars.usda.gov/Services/docs.htm?docid=13793#release

Wang, Ming-Dauh, et al. "Bayesian Modeling of Time Response and Dose Response for Predictive Interim Analysis of a Clinical Trial." New Developments in Statistical Modeling, Inference and Application. Springer International Publishing, 2016. 107-119.

Zhu, Wufu, et al. "Design, synthesis, and docking studies of phenylpicolinamide derivatives bearing 1H-pyrrolo [2, 3-b] pyridine moiety as c-Met inhibitors." Bioorganic & medicinal chemistry 24.4 (2016): 812-819.