Mechanistic Model-based Drug Oral Absorption Analysis

by

Kai Wang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Pharmaceutical Sciences)
in the University of Michigan
2021

Doctoral Committee:

Professor Gordon L. Amidon, Chair
Professor David E. Smith, Co-Chair
Adjunct Professor Sunny Chapel
Professor Kerby Shedden
Professor Duxin Sun

Kai Wang

kaiwangy@umich.edu

ORCID iD: 0000-0002-4475-4459

Acknowledgements

The deepest appreciation goes to my advisor, Dr. Gordon Amidon, one of the most influential pharmaceutical scientists in the world, for his guidance and inspiration in my adventure of gaining insight into drug oral absorption via quantitative models, for the flexibility and support he gave me in exploring all the possibilities in my career, and for his empathy, patience and encouragement during the days at graduate school. I also would like to thank Dr. David Smith, who helped me with the transition to graduate school study after I arrived at US and course selection especially in the first two years at University of Michigan. To my committee members, Dr. Duxin Sun, Dr. Kerby Shedden and Dr. Sunny Chapel, I feel very grateful for your time and comments offered throughout the process.

In addition, a huge thank you to Dr. Rose Feng, who helped me get more hands-on experience of developing population models. The side projects I worked with Dr. Feng gave me opportunities to get exposure to various fantastic models and communicate with wonderful pharmacometricians at conferences. Many thanks to Gail Benninghoff for her great work in coordinating my research in the lab and connection with my advisor, and Dr. Gregory Amidon for his help in 2020 to make my first department seminar possible.

Last but not least, I extend my gratitude to Jordan Peterson, Alfred Adler and some Chinese stand-up comedian celebrities for their work and thoughts to help me get through the dark moments, to my two newly adopted cats Lanlan (a Russian blue mix, super serious in eating and sleeping,

Table of Contents

List of Tables

List of Figures

Abstract

The thesis research consists of two projects: the major focus is using population approach to account for the systemic availability variability of Mycophenolate Mofetil (MMF) in human and providing insights into developing a predictive bioequivalence (BE) test. A separate small project on text mining via natural language processing (NLP) is included for drug Biopharmaceutics Classification System (BCS) classification.

The prodrug MMF, which is pre-systemically hydrolyzed into the pharmacologically active compound Mycophenolic Acid (MPA), has been widely used for the prophylaxis of acute allograft rejection in solid organ transplantation. However, the huge variability in plasma level makes MMF development difficult due to the great challenge of meeting the traditional BE limits. Numerous models have been developed in the past decade to explain the variability with the emphasis on characterizing the enterohepatic circulation (EHC), while the variability arising from absorption, can also contribute to the remarkable MPA variability to a great extent, but has been ignored for this BCS class 2 drug.

Two population pharmacokinetics (PK) models of MMF focusing on the absorption process were developed based on the plasma concentrations of MPA and its major metabolite MPAG in a long-term MMF treatment on liver transplant patients. The MPA PK profiles were best characterized by a two-compartment disposition model with zero inter-individual variability (IIV) of elimination

(K20), lag time (Tlag) but considerable inter-occasion variability (IOV) regarding systemic appearance rate (Ka), K20 and volume of distribution (V2). The second model took into consideration the EHC by including MPAG profiles as well. The results from both models showcased that the within-subject variability (WSV) of the MMF's systemic appearance played a much more significant role than the IIV. The large WSV can be mechanistically explained by the gastrointestinal (GI) physiological dynamics, especially gastric emptying (GE) in the fasted state regulated by migrating motor complex (MMC) and in the fed state by the caloric content with irregular patterns of GI motility and secretion. The results implied that dosing under fed conditions was recommended for the *in vivo* clinical BE study of MMF to reduce the WSV and that developing a predictive *in vitro* dissolution test with sufficient simulation of the GI physiological dynamics would be a good surrogate.

The second project explored the application of NLP in drug BCS classification. NLP, a confluency of artificial intelligence and computational linguistics, has gained widespread popularity in tech companies for machine translation, chatbot system, etc. In biotech and pharmaceutical industry, NLP-based text mining has been utilized to transform text information for decision support in multiple areas, including gene disease mapping, biomarker discovery, drug-drug interaction, and pharmacovigilance.

The BCS system, designed to recommend a waiver of *in vivo* bioavailability and BE studies for immediate-release (IR) solid oral drug products, classifies drugs based on their aqueous solubility under physiological pHs and intestinal permeability. However, there's no complete summarization of drug BCS classification information published to date. This project extracted solubility,

permeability and/or BCS class information of IR solid oral dosage forms in the human GI tract or simulated *in vitro* experiments from the FDA orange book, drug labels, FDA review documents, along with some selected literatures, and identified drug BCS class via NLP technology. The text mining results can be one of the key components in building up a database containing drug oral absorption information for drug discovery and development.

Chapter 1 Physiological Dynamics in Upper Gastrointestinal Tract and BCS Class 2 Drug Oral Absorption Models in Healthy Adult Human

1.1 Introduction

As the most prevalent drug delivery route, the upsides of oral dosing include easy administration, good patient compliance, and low manufacturing cost. However, one of biggest hurdles of developing an oral dosage form is the complexity of oral absorption process, which depends on GI tract environment under inter-digestive and/or post-prandial conditions, the physicochemical properties of the drug and the properties of dosage forms[1,2].

BCS was developed to aid in understanding and predicting drug oral absorption in the GI tract and has been endorsed by the regulatory agencies worldwide and the ICH for quality control in manufacture and marketing, as well as biowaiver of *in vivo* BE studies for IR dosage forms[3,4]. In this system, drug substances are classified into four classes on basis of their aqueous solubility in the GI physiological pH range (pH: 1.0-6.8) and their intestinal permeability (typically jejunal mucosal membrane):

- Class 1: high solubility, high permeability;

- Class 2: low solubility, high permeability;

- Class 3: high solubility, low permeability;

- Class 4: low solubility, high solubility[5].

According to U.S. FDA guidance, a drug substance is considered "*highly soluble*" when the highest strength is soluble in 250 mL or less of aqueous media within the pH range of 1-6.8 at 37±1°C, and "*high permeable*" when the systemic bioavailability or the extent of absorption in humans is determined to be 85 percent or more of an administered dose based on a mass balance determination (along with evidence showing stability of the drug in the GI tract) or in comparison to an intravenous reference dose. The definition of a high permeability drug was derived from a landmark publication of human jejunal permeability and the fraction drug absorbed[5]. A drug product is "*rapidly dissolving*" when a mean of 85 percent or more of the labeled amount of the drug substance dissolves within 30 minutes, using United States Pharmacopeia (USP) Apparatus 1 at 100 rpm or Apparatus 2 at 50 rpm (or at 75 rpm when appropriately justified in a volume of 500 mL or less (or 900 mL when appropriately justified) in each of the following media: (1) 0.1 N HCl or Simulated Gastric Fluid USP without enzymes; (2) a pH 4.5 buffer; and (3) a pH 6.8 buffer or Simulated Intestinal Fluid USP without enzymes[3].

The physiological dynamics in the upper GI tract (stomach, duodenum, and proximal jejunum) in different states (fasted or fed) is presumably to be the major source of variabilities in the dissolution and subsequent systemic availability ("absorption") profiles of drug products with active pharmaceutical ingredient (API) particularly having low solubility in the GI tract, especially BCS class 2 drugs[6]. In virtually all the cases, dissolution in the GI fluid milieu is a prerequisite for molecular drug absorption following oral administration. For BCS class 2 drugs, the dissolution process inevitably introduces considerable variability due to the vulnerability of the drug product to the environmental changes at absorptive sites, making it a determining factor in drug oral uptake and systemic appearance. The great variability observed in the systemic appearance of BCS class 2 orally dosed drugs is considered as one of the top concerns in the design of BE studies and also

in optimizing therapeutic drug monitoring (TDM). Therefore, a mechanistic characterization of the variability with regard to systemic appearance for this type of drug products from the perspective of dissolution regulated by GI physiological dynamics and a further application of this knowledge to the prediction of drug oral absorption process would be beneficial for BCS class 2 drug development and clinical efficacy and safety. Because of the high permeability nature, most of BCS class 2 drugs are absorbed in the duodenum and proximal jejunum after being emptied from stomach and dissolved in the GI fluid. This chapter is focused on reviewing the current studies regarding physiological dynamics in the upper GI tract and its relationship to drug oral absorption models in healthy adult humans.

1.2 Physiology of upper GI tract and organs draining to it

Before delving deep into the physiological dynamics affecting BCS class 2 drug oral uptake, it is necessary to discuss the upper GI tract under normal circumstances, so this section will center on the physiology of stomach, duodenum, jejunum, the pancreas and biliary system most relevant to drug oral absorption, as well as the underlying physiological basis.

1.2.1 Functional anatomy

1.2.1.1 Stomach
As the organ that initiates the absorption process in the gut, the stomach (Figure 1.1A) accommodates the ingesta upon its arrival in the fundus, gradually transfer the ingesta to the body and antrum, mixes it with gastric secretions, breaks it down to small-sized particles, and finally empty the luminal contents at a rate that allows efficient digestion and absorption in the more distal segments of the small intestine.

Functionally, the stomach can be divided into the proximal gastric reservoir and the distal gastric pump for considerations of motility[7] with the sphincter pylorus (antral musculature) at the gastroduodenal junction. The proximal stomach, consisting of the fundus, and proximal portion of the body, serves primarily as a reservoir and to move gastric contents to the distal stomach[8,9]. The gastric pump, including the distal portion of the body and the antrum, serves predominantly to grind and triturate ingesta and empty the stomach[9,10]. In addition, the pylorus is a narrow opening with folded, redundant mucosa. The periodic contractions generated at pylorus support for intermittent contractions. These properties enable it to control the amount and size (< 1 mm) of ingesta that can exit the stomach, making it a mechanical barrier and sieve. Conversely, when the pylorus opens and allows passage of luminal content to distend the duodenum, duodenal motor activity is stimulated[8-10]. Because the pylorus is an electric isolator and there exist different pacemakers in the stomach and duodenum, distinct frequencies between antral and duodenal contractions are observed. It is termed Antro-pyloro-duodenal coordination. In this coordination, the duodenum can contract three to four times during an antral wave. The contractions of the proximal duodenum cease during the phases of gastric emptying.[9]

The stomach can also be divided into two functional regions based on its secretory function: the oxyntic gland area and the pyloric gland area, respectively[11,12]. The oxyntic (oxys, Greek for acid) area, including the fundus and body of the stomach, contains approximately 75% of the gastric glands, so-called oxyntic glands. The characteristic secretions of the stomach are produced by specialized cell types in these glands. The most significant secretory product of the stomach, in the regard of drug oral absorption, is gastric acid (HCl). The stomach secretes around 400 mmol HCl per day, almost all of which enters the duodenum via gastric emptying[13], making it one of the crucial components of duodenal fluid. The concentration of H^+ reflected in the pH is typically 1-2

in the fasted state (0.01-0.1 M)[14]. The antrum of the stomach gains its name for being immediately proximal to the pylorus, contains glands that secrete gastrin, the primary regulator of post-prandial gastric secretion.

1.2.1.2 Duodenal cluster unit
Together with the pancreas and biliary system, the duodenum makes up the duodenal cluster unit. This segment of the gastrointestinal system plays a critical role in the regulation of digestion and absorption[10]. Figure1.1B depicts the overall anatomy of duodenum and its relevant organs.

Pancreas

The exocrine pancreas is the primary source of enzymes for digestion and bicarbonate for the GI predominant buffer. As alluded to above, large volume of strong acid HCl secreted by the stomach enters the duodenum with gastric emptying, placing the duodenal mucosa at risk for irreversible cellular acidification. In response to this acid onslaught, the duodenal cluster unit secretes bicarbonate to neutralize the HCl in the duodenal lumen. Quantitatively, the pancreas plays the major role among all the organs of the duodenal cluster unit in supplying the bicarbonate necessary to neutralize gastric[10].

The secretion of pancreatic bicarbonate is regulated by the secretin from the S cells in the duodenal mucosa. The S cells can be considered to act functionally as pH meters, sensing the acidity of the luminal contents (Figure 1.2). As the pH falls, due to the entry of gastric acid, secretin is released from the S cells and travels through the bloodstream to bind to receptors on pancreatic duct cells, as well as on epithelial cells lining the bile ducts and the duodenum itself. These cells, in turn, are evoked to secrete bicarbonate into the duodenal lumen, thus causing a rise in pH that will eventually shut off secretin release. The pancreas is quantitatively the most important in the

bicarbonate secretory response, although the ability of duodenal epithelial cells to secrete bicarbonate may be critically important to protect them from gastric acid, especially in the first part of the duodenum, which is proximal to the site of entry of the pancreatic juice and bile. The threshold for secretin release appears to be a luminal pH of less than 4.5.[10]

Biliary system

The biliary system, consisting of biliary tract and gallbladder, is responsible for transporting bile out of the liver into the GI lumen. The secretory products released by the biliary system include bile acids, phospholipids, cholesterol, bilirubin, and bicarbonate. Among them bile acids aid in drug solubility and dissolution by forming mixed micelles, when above their critical micelle concentration (CMC).[10] Biliary secretions mix with those coming from the pancreas at the common bile duct, and flow into the duodenal lumen in a controlled fashion when the sphincter relaxes in response to neurohumoral influences.[10]

Duodenum

The absorptive capacity of the small intestine (Figure 1.1C) is remarkably amplified attributed to its massively amplified surface area compared to that of stomach. As we know, the intestine is not a simple cylinder, but instead is amplified first by folds in the mucosa, then by the presence of crypts and villi, and finally by the presence of abundant microvilli on the apical poles of individual epithelial cells, increasing the overall surface area by a factor of 600-fold. This amplification of the surface area makes it the perfect place in the gut to handle luminal contents including drugs and nutrients.

The first segment of the small intestine, approximately 12 inch in length, is referred to as the duodenum, and begins as a bulb-shaped structure immediately distal to the pylorus. The motor

function of duodenum is to mix ingesta with GI secretions so that the luminal contents can be absorbed or propel ingesta relatively quickly so that it does not act as an obstruction to further gastric emptying. Duodenal mucosa also contributes to the secretion of bicarbonate, but its role is minor compared to pancreatic ductular cells.[10]

1.2.1.3 Jejunum

The jejunum serves as main absorptive site for most of nutrients and drugs in the healthy individual and has a markedly amplified surface area due to the presence of surface folds, tall and slender villi, and microvilli mentioned above. This increase in surface area, combined with cyclical discontinuous contractions caused by GI motility patterns (summarized in 1.2.2.1), lead to rapid absorption of nutrients and highly permeable drugs.
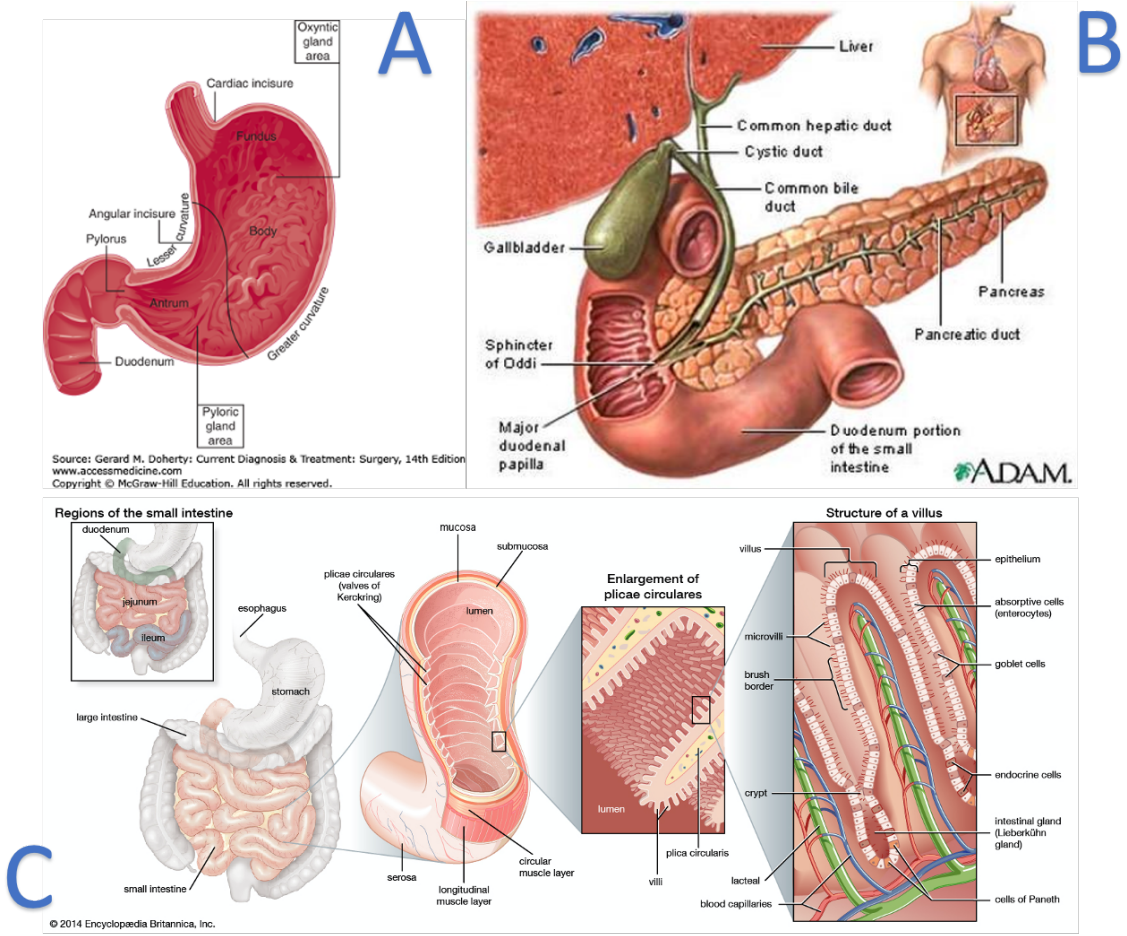
Figure 1.1 The overall anatomy of the stomach (A) [15], duodenal cluster unit (B) [16] and small intestine (C) [17].
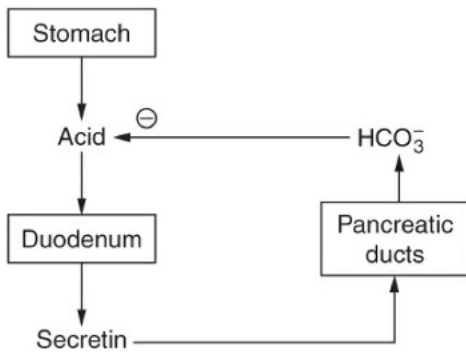
Figure 1.2 Regulation of duodenal pH by pancreatic bicarbonate secretion in response to gastric acid emptying.

1.2.2 Physiological basis

1.2.2.1 Motility

The mixing and propulsive motor functions of the gut to facilitate absorption are primarily governed by cyclical phasic contractions. The rate of propulsion generally slows distally in the small intestine, while mixing intensifies[18].

In the inter-digestive state, phasic contractions become organized as a cyclical motor pattern migrating motor complex (MMC) typically initiated in the stomach or duodenum. This contractile activity consists of:

- Phase 1 activity that has little or no contractile activity;

- Phase 2 activity that shows an increasing number of intermittent but irregular and rarely propulsive contractions;

- Phase 3 activity that is a group of the largest amplitude peristaltic waves occurring at their maximum frequency, and the entire group of contractions migrates distally over a long distance, usually the entire small intestine, in an organized fashion (Figure 1.3)[8-10,19]. In the absence of feeding, MMC cycle lasts around 100 mins: the quiescent phase 1 comprises 40-60% of the cycle, phase 2 comprises 20-30% of the overall cycle, and the phase 3 involves a 5-10 min burst of intense phasic contractions. The onset of phase 3 is correlated with plasma motilin level[10]. It is noteworthy that the pylorus opens fully in the phase 3 of MMC and the actual propulsion of luminal contents is caused by the propagating phasic contractions within phase 3 activity. As the contractile activity propagates it becomes less spatiotemporally organized resulting in slower propulsion rates in the distal small bowel.[10,19]
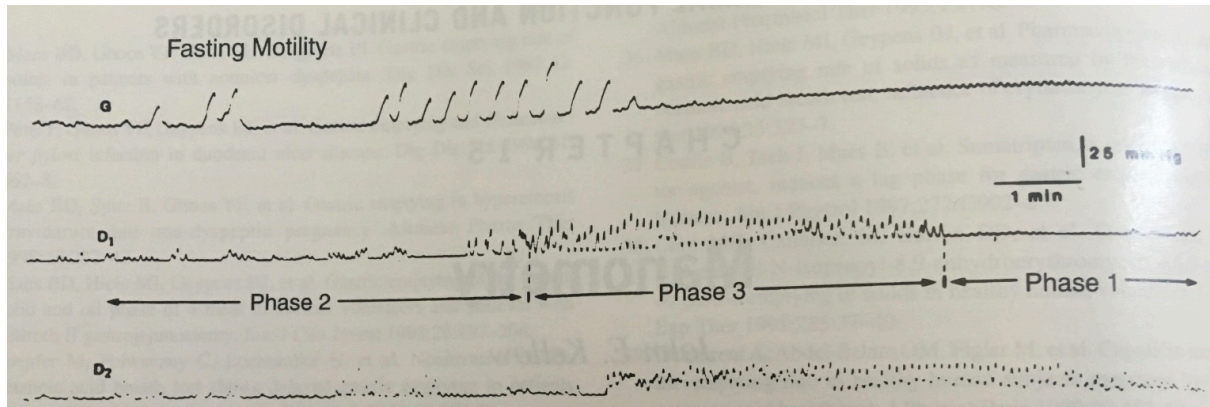
Figure 1.3 The three phases of MMC in the fasted state recorded using a stationary water-perfused catheter system at gastric antrum (G) and the first and second parts of the duodenum (D1 and D2, respectively).

Following ingestion of a meal or any type of nutrients, there is an immediate change in the contractile pattern of the stomach and small intestine to the fed pattern. The pylorus remains closed for prolonged periods with only intermittent openings that allow small particles (< 1-2 mm) to enter the duodenum. The motility fed pattern of the small intestine differs from the fasted pattern with variable frequency, amplitude, and propagation to enable the mixing of luminal contents and their subsequent aboral transport. Up to 50% of the phasic contractions in the duodenum is reported to actually move the luminal contents orally[10]. The duration of the fed pattern of motility in the small intestine depends on the caloric content of the meal ingested as well as its composition[19].

1.2.2.2 Secretion

Ample experimental data support a coupling between inter-digestive GI motility and secretion[20-22]. The MMC also cycles in phase with contractile activity of the gallbladder and relaxation of the sphincter of Oddi, as well as with periodic increases in secretory function of the intestine and organs draining into it (Table 1.1)[10]. Phase 1 secretion is characterized by minimal exocrine pancreatic and gastric secretion, absence of bile in the duodenum, and maybe a refractory phase for pancreatic secretion. During phase 2 secretion increases. As the irregular motor activity

10

intensifies, exocrine pancreatic secretion into the duodenum reaches maximal inter-digestive secretory rates several minutes before the onset of phase 3[23]. During phase 3, the exocrine pancreatic secretion decreases until the upper GI tract resumes the quiescence of phase 1. Bile acid secretion parallels pancreatic enzyme secretion and peaks during late phase 2[24]. Bile entry into the duodenum is preceded by gallbladder contraction and emptying of around 25% of gallbladder contents, which accounts for most of the bile within the duodenum during phase 2. In contrast, peak gastric acid secretion and bicarbonate secretion into the duodenum occur slightly later coinciding with the start of phase 3.[25]

Table 1.1 Secretion and absorption of the gut marked by MMC

	Components	Phase 1	Late phase 2	Early Phase 3
Secretion	HCl	min		max
	HCO3-	min		max
	Bile	min	max	
	Pancreatic enzyme	min	max	
Absorption	Water	max	min	

Under the post-prandial, GI secretion can generally be divided into three phases[10]:

- Cephalic Phase: in response to any sight, smell, taste and even the thought of the food, the vagal outflow initiates gastric secretory behavior to readies the stomach to receive the meal and 20-25% of pancreatic secretion with high concentration of digestive enzyme;

- Gastric Phase: in addition to vagal influences continuing from the cephalic phase, gastric secretion is amplified further by mechanical and chemical stimuli when the stomach distends to accommodate the volume of the meal. These vago-vagal reflexes also transmit information downstream to ready more distal segments of the intestine to receive the meal. The combined influence of neurocrine and endocrine signals makes this phase quantitatively most important for gastric secretion. At this phase, the pancreatic secretion is only 10% also with high concentration of enzymes;

- Intestinal Phase: as the meal moves out of the stomach into the duodenum, the gastric secretion is suppressed in response to the fall of pH in the lumen, while high volumes (60-70%) of pancreatic juice is produced with decreased concentrations of protein.

1.3 Physiological dynamics in the upper GI tract and the corresponding models

In a typical BE study, solid orally administrated drug products like tablets or capsules are usually given with a glass of water (250 mL) in the normal state. We'll see below how variable this "normal" GI fasted state is due to the MMC described above and further how different the fasted and fed state are in terms of motility and secretion physiologically.

1.3.1 Gastric emptying (GE) and upper small intestinal transit variations
The rate of GE and upper small intestinal transit is negatively associated with the extent of absorption: if motility is increased, hastening the transit of substances along the intestine, there will be less time for absorption to take place; if transit is slowed, absorption can catch up with the presented the volume luminal contents.

1.3.1.1 Inter-digestive
Conventionally, human gastric emptying is regarded as continuous first-order process, however, an increasing amount of research demonstrated the process of gastric emptying to be far from that simple. For liquid volumes of 240 to 800 mL, experimental measurements of gastric emptying half-time varied from 8 to 18 minutes[26-29]. In the fasted state, the gastric emptying rate of liquid was shown to heavily rely on MMC phases and the larger volume of 200 mL phenol red solution leaves the stomach with a half-life of 11.8 min, being less dependent on gastric motility than the smaller volume of 50 ml solution[30]. In the study of Mudie et al., 75% of the subjects displayed

first-order emptying patterns while 25% had non-first order, biphasic emptying after drinking 240 mL water [31].

Likewise, human intestinal transit is traditionally modeled with continuous first-order approximation. There is very limited published experimental evidence regarding intestinal transit in the fasted state. But according to the results of mass transport analysis of phenol red, first-order process is obviously inadequate to characterize duodenal and jejunal transit[32]. Considering the high frequency and incidence of MMC in the proximal intestine, MMC presumably plays a critical role in the complicated hydrodynamics pattern in proximal small intestine.

GE and intestinal transit are more likely to occur in a discontinuous fashion inferred from experimental evidence and GI physiology theory. After Schiller et al. reported that that fluid in the fasted small intestine is not a continuous watery compartment but rather in discrete packets of varying volumes, Mudie et al. further quantified the volume and number of water pockets in the small intestine of fasted healthy humans.[31,33] The resting small bowel water was distributed in 8 ± 1 pockets of 4 ± 1 mL on average each, rose to 15 ± 1 pockets of 6 ± 2 mL each at peak time, and 16 ± 3 pockets of 5 ± 1 mL each at 45 min before gradually returning to the baseline level. Regardless of water intake, the number of liquid pockets in the smaller 0.5-2.5 mL size bin was markedly higher than all the larger bin sizes and that there was no significant difference between the remaining larger bin sizes, but most of the total volume of liquid was found in the larger pockets. At the time point of 45 min after water administration, the smallest pockets (0.5-2.5 mL) account for less than 5% of the total liquid volume, the smaller pockets (2.5-20 mL) claims about 40% of the total volume, and the largest amount of liquid (~60% of the volume) is contained in a

small number of large pockets (> 20 mL)[31]. According to the overview of GI physiology described above, it's natural to relate the discrete fluid packets emptied from the stomach with time-varying volumes phenomena to the varying onset, frequency, amplitude, duration, direction and distance of antral contraction and small intestinal peristalsis, the phasic secretion of the stomach, pancreas, biliary system and duodenal mucosal, and the different frequency and extent of opening of pylorus and Sphicter of Oddi during MMC cycle, thus making any detailed mechanistic model of this process extremely complex.

The identification of MMC-driven discontinuous, small volume of fluid packets helps to explain the variability of oral absorption of BCS class 2 drug and consequently has significant implications for TDM and the design of BE studies of this type of drug products. To begin with, the randomness of the dosing time relative to MMC phases undoubtably brings in the variation of the delivery of drug to the absorptive site. If the drug is administrated at the beginning of MMC phase 1, then it will likely be kept in the stomach for a significant period of time by the pylorus; To the contrary, if the drug is dosed during phase 3, the strong contractile period, then it will pass the stomach through the fully open pylorus into the small intestine without residence. In this case, the absorption profiles would also be quite variable due to distinct GI motility patterns. In addition, the presence of the small fluid packets would contribute to the within-subject variability of drug Cmax and Tmax. If the volume of a fluid packet is lower than the volume needed to completely dissolve the drug particles it encloses, the undissolved drug will not be transported across the GI mucosal cells and show up in the circulating system together with the dissolved counterpart. These drug particles are often held back in the GI lumen until they encounter appropriate fluid packets. For instance, in a study of Schiller et al, only 50% of the ingested non-disintegrating capsules were

surrounded by fluid in the fasted small intestine[33]. Hence, the randomness of the timing and size of fluid packet drug particles are exposed to may lead to the uncertainty of the systemic appearance of low solubility drugs, especially the amplitude of Cmax and the timing of Tmax. Furthermore, the discontinuity of the fluid packet would inevitably incur higher variability than a continuous fluid flow system. If a drug product is completely dissolved in the stomach, it will remain there until the arrival of the next antral contraction; if a drug product is emptied from the stomach as the undissolved particles, none of them will be dissolved in quiescent period before the next wave of intestinal peristalsis brings fluid packets. So we speculate there are jumps in the cumulative plot of dissolved drug vs time. Collectively, huge intra-individual variability would be anticipated in the dissolution profiles and systemic appearance due to this super dynamic GI transit feature. CellCept, the IR solid oral dosage form of MMF, is one example of BCS class 2 under TDM. The major variability of its within-subject variability comes from GE, which will be discussed in detail in Chapter 2. The inter-individual variability of its systemic appearance will be captured by sufficiently simulating the GE dynamics, therefore the Cmax will be predicted more accurately in TDM and also help the BE study to focus on drug product difference by minimizing the confounding effect of GI dynamics.

1.3.1.2 Post-prandial
Food ingestion is known to profoundly alter the GI environment by interrupting the cyclical MMC. As discussed previously, the motor patterned of the antrum under fed conditions are featured as mixing and grinding, while in small intestine they are called mixing and segmentation[10]. With propulsion occurring significantly less aborally and more orally, a delayed GE and prolonged intestinal transit can be expected[34].

It has been demonstrated that the caloric load and the type of nutrients ingested play a critical role in regulating the motor response the upper GI tract (Figure 1.4). Fed pattern durations of about 180, 360, and 410 mins were reported to occur after meals of 630, 1260, and 2520 kJ, respectively[35]. In general, fat produces a longer fed pattern than carbohydrate or protein, but the timing of the meal does not appear to affect the GI transit for the same nutrient load. The GE of inert liquids is most rapidly, while for the liquid contains nutrients, a rapid initial phase is then followed by slowed exit, apparently reflecting feedback from the small intestine. Moreover, emptying of solids from the stomach is slower yet, with a half time of approximately 1–2 hours during which retropulsion and mixing take place. This lag phase is increased in duration if large particles are swallowed as a whole. After the lag phase, a linear phase of emptying of a particulate suspension occurs at a relatively constant rate, with the size of 1-2 mm and the volume of 2-3 mL. Likewise, intestinal transit is slowed in proportional to the calory and also depends on the type of nutrients.[10,34]
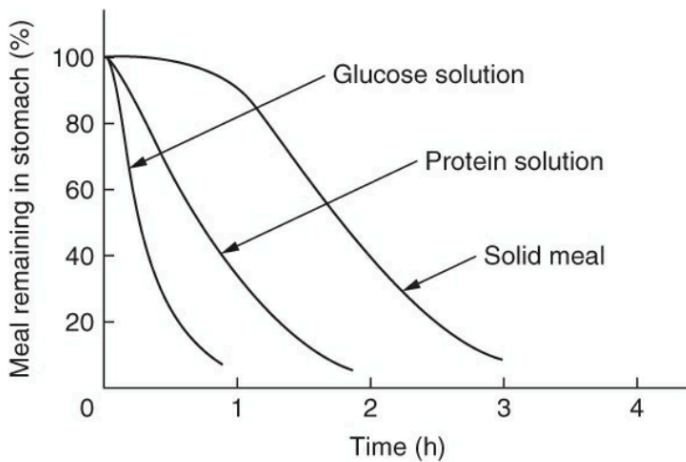
Figure 1.4 GE time curve of different types of gastric contents

Therefore, the regulatory agency like FDA published a standard for the food-effect bioavailability (BA) and fed BE studies. A high-fat (approximately 50 percent of total caloric content of the meal) and high-calorie (approximately 800 to 1000 calories) meal is recommended as a test meal for food-effect BA and fed BE studies. This test meal should derive approximately 150, 250, and 500-600 calories from protein, carbohydrate, and fat, respectively. In NDAs, it is recognized that a sponsor can choose to conduct food-effect BA studies using meals with different combinations of fats, carbohydrates, and proteins for exploratory or label purposes. However, one of the meals for the food-effect BA studies should be the high-fat, high-calorie test meal described above. As for the fed treatment, following an overnight fast of at least 10 hours, subjects should start the recommended meal 30 minutes prior to administration of the drug product. Study subjects should eat this meal in 30 minutes or less; however, the drug product should be administered 30 minutes after start of the meal. The drug product should be administered with 240 mL (8 fluid ounces) of water. No food should be allowed for at least 4 hours post-dose. Water can be allowed as desired except for one hour before and after drug administration. Subjects should receive standardized meals scheduled at the same time in each period of the study.[36]

1.3.1.3 Models

A great many of endeavors have been made to model drug oral absorption, however, the modeling and simulation of BCS class 2 drug systemic appearance is still tough partly due to the variable GI transit. How the GI transit kinetics are incorporated in published models are included in Table 1.2.

A typical empirical absorption model usually assumes a well-mixed GI lumen compartment with zero-order or first-order constant absorption coefficient with or without a lag time. Another commonly used empirical model is the one including two parallel first-order constant absorption

17

coefficient with two lag times (Model 1) to describe double peaks in the absorption phase. Sometimes, a simultaneous or a sequential combination of zero-order and first-order absorption is employed, for example the first-order followed by zero-order absorption in the case of Model 2. If the zero-order and/or first-order absorption assumption is still inadequately to describe the complicated absorption profile, Weibull function and its variants may worth a try. The idea of Weibull function has been used in both modeling (Model 3) and simulation in fasted (Model 4) and fed state (Model 5) to account for the GI motility variability. Occasionally, the Michaelis-Menten (MM) equation can be introduced to describe the saturable absorption phenomenon (Model 6). The TPLAG model (Model 7) simulated the oral absorption with two first-order constant absorption coefficients with two lag times conceptually, to account for the dynamics in fluid absorption or secretion, absorption surface area, and time-varying motility caused by two phases of GE. More sophisticated models include the one using a sum of inverse Gaussian functions (Model 8) and the one appling the law of Laplace, Hooke, and Poisseuille to derive a nonlinear first order deterministic elimination GE model (Model 9) in fasted and fed state, respectively. These empirical modelings provide estimates of the population mean of absorption kinetics implying GI transit and the corresponding variability, which can be used as the prior information or reference value for the subsequent prediction model, but nothing related with the GI physiological mechanism. The empirical simulations, though too simple, are the exploratory works toward building physiologically-based absorption models.

The mechanistic models handle the oral absorption from a bottom-up approach. The parameters of GI transit are explicitly specified in the model based on experimental data or literature search rather than inferred from the parameter estimation of the absorption kinetics in the empirical modelings.

At the early phase of developing such oral absorption simulation models, the GI transit were just assumed to be a zero-order or first-order process with constant transit coefficients (Model 10-12). A higher transit rate was applied in the fasted state and a lower transit rate in the fed state. Then more physiological information was introduced later, for example unlike the CAT model that supposes all the small intestinal compartment share the same transit coefficient, the refined ACAT model (Model 13) included compartment-dependent physiological parameters to represent the distinct features of each segment of human GI tract. Model 14 assumed first-order GE with three constant coefficients representing the three phases of MMC, but in the real work, the change of motility pattern is not a jump step. Model 15 designed another bypass compartment to accommodate the huge variability resulting from the dynamic GI motility patter and explored the potential relationship of the fraction of fluid flowing to the bypass compartment vs time to MMC phase 3 post-dose. Model 16 started to simulate MMC pattern directly in a non-linear way with the Fourier approximation for GE rate and a sigmoidal decay function for the lag time. The absorption profiles could also be modeled by a delayed differential equation (DDE) and a sine function to fit the complicated multiple peaks in the absorption phase. Interestingly, Model 17 employed PCA to investigate the impact of GE on individual PK parameters, trying to link the empirical modeling work with GI physiology. Considering the complexity of developing a discrete model, most of the oral absorption models handle the discontinuous GI transit continuously and smoothly with lag time(s) from drug administration to systemic appearance, while Model 18-20 were stochastic models designed to capture these spurts in a discrete system. Meta-model was used to modeling the probability of flow from one compartment to another by first-order transit in the fasted and fed state (Model 18). Due to the irregular pattern of pylorus opening in the fed state, the GE process was simulated with the power exponential function of standard Wiener process

representing the irregular decrements of gastric contents after a meal in Model 19. Model 20 simulated the occurrence of GE with nonhomogeneous Poisson process and the intensity of GE in the form of the volume of fluid pulse packets to account for time-varying contractions during three MMC phases in the fasted state. Obviously, the mechanistic simulation models not only allow for complicated model structure, but also enable prediction of GI luminal contents and systemic appearance on the basis of GI physiological dynamics. More examples of how to model the impact of GI transit on drug oral absorption can be found in Ahmad et al's work on a quantitative review and meta-models in the fasted and fed state[37].

Table 1.2 A summary of GI dynamic variables in published models

No	Model description	Reference
1	A population PK model (fasted): two parallel first-order absorption: one represents the immediate first-order absorption from stomach and upper intestine and the other represents the delayed first-order absorption process from the lower intestine	38
2	A population PK model (fasted): First-order followed by zero-order absorption	39
3	A population PK model (fasted): Two Weibull function absorption	40
4	A GE simulation model (fasted): The sum of two Weibull models in different proportions to account for the motility variability of MMC	41
5	A GE simulation model (fed): A modified Weibull model to represent time-varying gastric retention	42
6	A population PK model (fasted): A saturable absorption model (Michaelis-Menten equation)	43
7	A two-phase simulation model with lag time (TPLAG, fasted): 1) First-order absorption with time-dependent absorption coefficient, representing two phases of GE; 2) Constant absorption coefficient in each of the two phases of GE;	44
8	A population PK model (fasted): A parametric deconvolution method using a sum of inverse Gaussian functions to describe the absorption profiles	45
9	A GE simulation model (fed): 1) Applied the law of Laplace, Hooke, and Poisseuille to derive a nonlinear first order deterministic elimination GE model assuming the initial volume of the stomach is greater than 300 mL; 2) If the initial volume of the stomach is less than 300 mL, some correction is needed	46
10	A two-tank perfect-mixing tank simulation model: 1) First-order absorption with constant coefficients; 2) First-order transit with constant coefficients—assume higher Q in the fasted state and lower Q in the fed state; 3) Different constant intestinal pH among different populations	47
11	A four-compartment simulation model: 1) Zero-order and first-order GE with different constant coefficients with or without lag time to represent GE in fasted and fed state; 2) Bile salt input as a bolus in fed state and zero in fasted state;	48

	3) First-order transit and absorption with constant coefficients	

12	The Compartmental Absorption and Transit (CAT)—a nine-compartment simulation model (fasted and fed):
	   1) First order GE and intestinal transit with constant coefficients;
	   2) The seven small intestinal compartments have equal transit time;
	   3) First-order absorption with constant coefficients;
		49

13	The Advanced Compartmental Absorption and Transit (ACAT)-- a 18- compartment simulation model (fasted and fed):
	   1) First order GE and intestinal transit with constant coefficients;
	   2) Include compartment-dependent physiological parameter (e.g., pH, transit rate coefficients);
		50

14	A one-compartment simulation model (fasted): Three distinct constant absorption coefficients in the three phases of MMC;	51

15	A five-compartment simulation model:
	   1) Bypass compartment to facilitate transport directly to the duodenum in the fasted state;
	   2) Bypass compartment to facilitate transport directly to the antrum of the stomach in the fed state;
	   3) First-order transit;
	   4) First-order absorption and zero-order secretion in duodenum and jejunum;
	   5) Negligible gastric secretion
		32

16	A eight-compartment simulation model (fasted):
	   1) Assumes GE is MMC-driven motility-dependent
	   2) GE rate with a Fourier series approximation;
	   3) Lag time with a sigmoidal decay function;
		52

17	Two population PK models (fasted):
	   1) The first PK model using Delay differential equations;
	   2) The second PK model using a sine function;
	   3) The impact of GE on PK parameters investigated by applying principle component analysis (PCA) to the individual parameter estimates
		53

18	A stochastic continuous-time Markov-chain meta-model (fasted and fed):
	   1) First-order transit represents the probability of flow from one compartment to another;
	   2) Two-compartment stomach + five-compartment small intestine with equal transit rate coefficient + four-compartment colon
		54

19	A stochastic GE simulation model (fed): GE with Weiner process	55

20	A stochastic GE with fluid packet simulation model (fasted):
	   1) the occurrence of GE with nonhomogeneous Poisson process to account for time-varying contractions during three MMC phases;
	   2) the volumetric effect in the stomach;
	   3) the volume of fluid packet is a function of the current stomach volume and mean packet size plus some deviations
		56

1.3.2 GI luminal pH fluctuations

For an ionizable low solubility drug, the solubility and dissolution rate would be greatly influenced by pH fluctuations, which is prominent in the upper GI tract with low buffer capacity. It is important to note that the buffer capacity comes from the fluid secretion into the GI lumen rather than reside in the GI fluid due to the unique chemistry of HCO_3^- and CO_2. Hence, characterization of the pH fluctuations would promote the prediction of the fraction of BCS class 2 drug dissolved in the GI lumen.

1.3.2.1 Inter-digestive

In fasted state, the stomach environment ranges from pH 2 to pH 4 in healthy individuals[57]. As the conduit of random gastric emptying in a random discrete fluid packet fashion[31] and motility-related pancreatic and biliary draining[25] in the low buffer capacity setting, the proximal small intestine experiences the most dramatic pH fluctuation throughout the whole small intestine and tapers down distally. A large range of pH fluctuation (pH: 1.71-7.57) in the proximal small intestine has been reported in Bart et al's work (Figure 1.5)[57]. For example, the duodenal pH is a dynamic physiological parameter ranging from 2.4 to 7.5[58] with considerable intra- and inter- individual variability (Figure 1.6)[59]. The non-uniform input rate of main strong acid HCl from gastric emptying (~17 mmol/h)[60] and main weak base HCO_3^- (~8 mmol/h)[61] from the duodenal cluster unit secretion, and limited buffer capacity of the bicarbonate/CO_2 buffer in human small intestinal fluids (mean: 2.26 mM/pH, range: 0.26-6.32 mM/pH)[57,62,63] could all be the culprits of pH fluctuation in the duodenum.
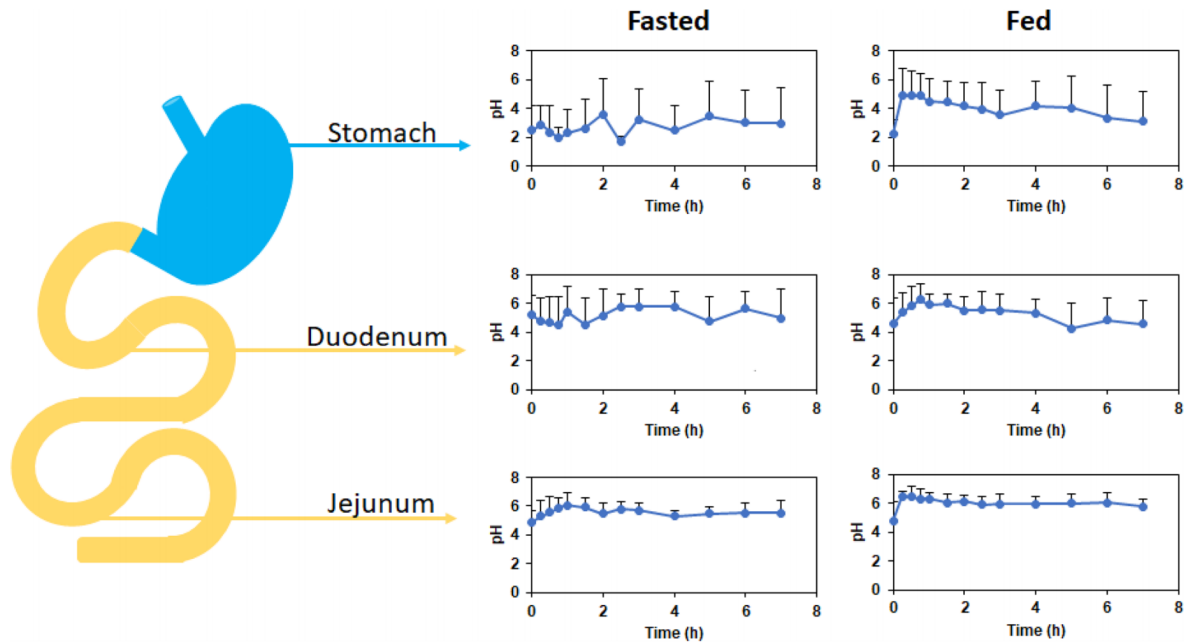
Figure 1.5 Mean pH versus time profiles in fasting (n = 20) and fed state (n = 17) conditions as measured in the stomach, the duodenum, and the jejunum (mean + SD). Data obtained from Koenigsknecht et al.[64]
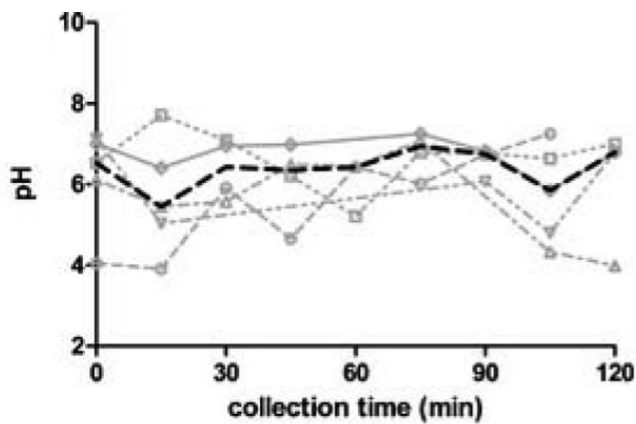
Figure 1.6 Individual pH versus time in fasting human duodenal fluid for five healthy subjects

1.3.2.2 Post-prandial

As shown in Figure 1.5, after food consumption, the gastric pH transiently increases, while the pH

levels in the duodenum and Jejunum do not fluctuate as dramatically as the gastric pH, and remain

around pH 4 to pH 6. The pH profile along the intestinal tract is dynamic under fed conditions and

largely depends on several factors, including the composition of food ingested, gastric secretion, and bile salt concentrations.[65]

1.3.2.3 Models

There are not many models in terms of the luminal pH fluctuation developed to date. Among the different types of oral absorption models in Table 1.2, only Model 10 and Model 13 (ACAT) take the effect of pH fluctuation into consideration. Meta-models of the impact of GI pH on oral drug absorption were summarized in Ahmad et al's work.[66]

1.4 Conclusion and discussion

The review revealed the relationship of GI physiological dynamics with the variability of BCS class 2 drug oral absorption via an overview of GI physiological knowledge and an update of endeavors that modelers have made so far to capture the GI dynamics, especially GI transit and luminal pH fluctuations. This would shed light on further accounting for high within-subject variability of the systemic appearance of BCS class drug 2, and provide insights into dosing recommendation and the design of BE studies.

However, due to the complexity of oral absorption process, it may be a good idea to switch from the traditional parametric models to some non-parametric learning, for example, the machine learning (ML) and artificial intelligent (AL) algorithm. After grinded into small particles, many high permeable drugs are emptied from the stomach to the duodenum in a random discrete fluid packet[31], mixed well with the GI secretion and then be absorbed at the duodenum or upper jejunum after the delivery of intestinal transit to the absorptive site. Modeling of the discontinuous fluid packet emptying and GI transit, the unstable GI secretion, GI fluid buffer capacity, and the impact

of the drug on local pH is never an easy task, ML/AL algorithm would potentially worth a try by avoiding proposing a specific model structure to handle all these complicated GI events. For example, dimensionality reduction could be used to help figure out where the variability of systemic appearance mainly comes from among all these GI events, by feeding the model with a pool of the BCS class 2 drug product absorption data. In addition, reinforcement learning (RL) is another fantastic strategy for predicting drug systemic appearance in the longitudinal PK data.

Chapter 2 Population Pharmacokinetics of Mycophenolate Mofetil in Liver Transplanted Recipients and its Implication for Developing a Predictive Bioequivalence Test

2.1 Introduction

Mycophenolate mofetil (MMF), initially launched into the market as CellCept® in 1995, has been widely used as an immunosuppressant for the prophylaxis of organ rejection after kidney, heart, or liver transplantation with 33[67] bioequivalent drug products to date[68].
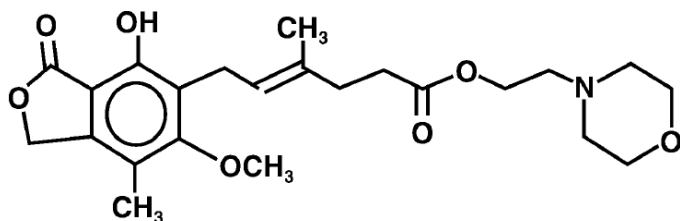
Figure 2.1 The chemical structure of MMF[68].

MMF has an empirical formular of $C_{23}H_{31}NO_7$ with a molecular weight of 433.50. The chemical structure of MMF is depicted in Figure 2.1. It is slighted soluble in water (43 ug/mL at pH 7.4) and soluble in acidic medium (4.27 mg/mL at pH 3.6). The pKa values for MMF are 5.6 for the morpholino group and 8.5 for the phenolic group. MMF is classified as class 2b drug under BCS for the low solubility in the physiological pH range, which is indicated in the theoretical solubility profile in Figure 2.2 relative to the highest approved formulation unit dose of 500 mg[67], and also the high permeability with an apparent partition coefficient in 1-octanol/water (pH 7.4) buffer solution of 238[68].
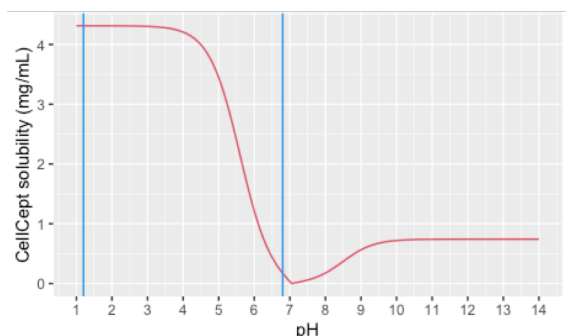
Figure 2.2 Theoretical solubility profile of CellCept® at pH 1.0-14.0. It was generated based on Henderson–Hasselbalch equation and isoelectric point. The pH range between the two blue lines represents the typical physiological pH range, 1.2-6.8.

The absorption process of MMF is complex, exhibiting varying plasma concentration-time profiles for the pre-systemically hydrolyzed active metabolite mycophenolic acid (MPA), lag time (Tlag), maximal concentration (Cmax), time to maximal concentration (Tmax), as well as double peaks at the absorption and post-absorption phases[69,70]. The systemic appearance, biotransformation, transport, bile secretion and gut microbiota metabolism and reabsorption processes are shown in Figure 2.3. Considering the low solubility, high permeability nature of a BCS class 2 drug, the systemic appearance of MMF as MPA can be significantly affected by the changes in the subjects' GI tract environment i.e. different motility states in fasted condition and fasted vs fed conditions. It has been demonstrated that the second MPA peak at the post-absorption phase is usually observed 6-12 hours post-dose and corresponds well to food intake times due to enterohepatic circulation (EHC)[68]. Therefore, the major metabolite of MPA, 7-O-MPA-β-glucuronide (MPAG), which plays a crucial role in EHC, should also be quantified, despite being pharmacologically inactive. Variations of any factors in the MPA/MPAG EHC can impact the variability of the second peak. The numerous potential sources of remarkable MPA PK variability are summarized in Figure 2.4.
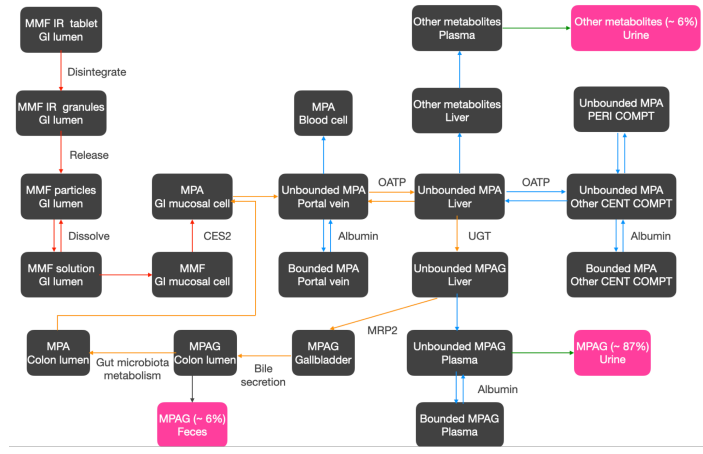
Figure 2.3 Overall outline of MMF PK in human body. The red, yellow, blue and green arrows indicate the process of systemic appearance, enterohepatic circulation (EHC), disposition, and excretion, respectively. The MMF immediate release drug product undergoes rapid disintegration and release process into the gastrointestinal (GI) tract after oral administration before fully dissolved in the acidic gastric fluid. Then it permeates the upper GI mucosal cell membrane and is pre-systemically hydrolyzed into the pharmacologically active compound mycophenolic acid (MPA) by carboxylesterase 2 (CES2)[71]. At clinically relevant concentration, MPA is 97% bound to plasma albumin. The unbounded MPA is uptaken into the liver primarily by OATP before converted into the 7-O-MPA-b-glucuronide (MPAG) and other metabolites mainly by UGT1A9 and UGT2B7[72,73]. MPAG is then actively transported to the gallbladder by MRP2[74], released to duodenum with bile secretion for the conversion back to MPA in the gut by microflora. With regard to the overall mass balance, orally administered radiolabeled MMF resulted in ~87% recovery of the administrated dose as MPAG, ~6% recovered as other metabolites in the urine and 6% recovered in feces. Negligible amount of drug is excreted as MPA (less than 1% of dose) in the urine[68].
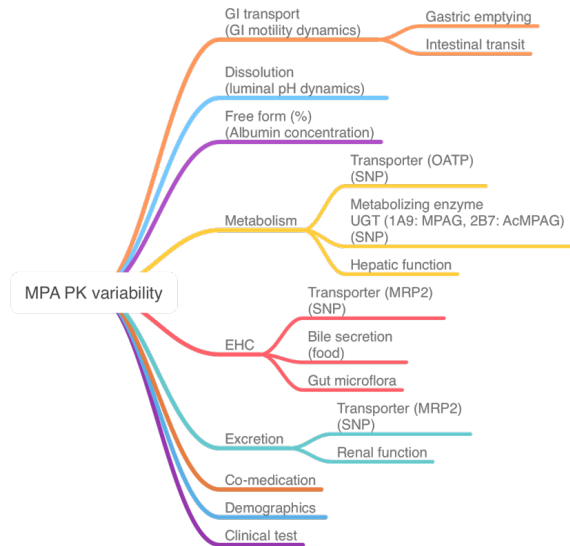
Figure 2.4 Potential factors contributing to MPA PK variability.

Like many other highly variable (HV) drugs, whose within-subject variability (%CV) in bioequivalence (BE) metric(s) Cmax and/or AUC >= 30% by definition[75,76], the 90% confidence

28

interval (CI) around the geometric mean ratio (GMR) of the test MMF/reference MMF of Cmax and AUC would inevitably exceed the BE limits (80%-125%)[77] with 18-24 (a common number of subject enrollment) subjects under two one-sided tests (TOST)[78] procedure as shown in Figure 2.5. As the width of the 90% CI depends upon the number of subjects in the study and the variability of the BE metrics, it's natural to either increase the number of subjects enrolled for the BE (as shown in Table 2.1) or adjust the BE limits using Reference-Scaled Average Bioequivalence (RSAB, $-\ln(1.25)\ \sigma_{WR}/\sigma_{W0} <= (\mu_T-\mu_R) <= \ln(1.25)\ \sigma_{WR}/\sigma_{W0},$ where σ^2_{WR} and σ^2_{W0} are the population WSV of the reference formulation and a pre-determined constant set by the regulatory agency, respectively) instead of the traditional Average Bioequivalence (ABE, $-\ln(1.25) <= (\mu_T-\mu_R) <= \ln(1.25)$), where μ_T and μ_R is the population average response of the log-transformed measure for the test formulation and reference formulation, respectively) based on the WSV[79]. For example, Patel et al[80] and Almeida et al[81] recruited 126 and 103 healthy subjects for BE study under fasting conditions, respectively, due to the high WSV of MMF. However, increasing the number of subjects is time-consuming and cost-expensive, and there are still some controversies in terms of the selection of the value of σ_{W0}, a better approach that can tell product difference between test and reference formulation of MMF or even the HV drugs without the confounding effect of WSV would potentially be a new BE test standard.
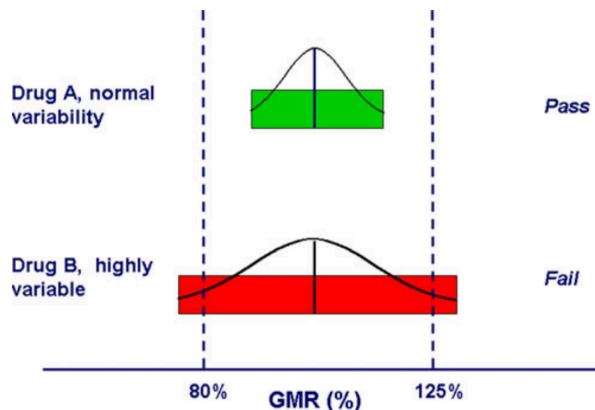
Figure 2.5 The 80–125% BE limits are represented along the x-axis as two "goal posts." The BE limits are compared to the hypothetical 90% CIs of the test/reference BE measure GMRs for two drugs, a drug with normal variability (Drug A) and an HV drug (Drug B). The 90% CIs of the two drugs are represented by colored bars. For drug A (normal variability), the 90% CI (green bar) meets the BE limits. For drug B (HV), the 90% CI (red bar) fails to meet the acceptance limits. As the width of the CI is influenced by the number of study subjects, in the hypothetical case of drug B, it is likely that the study would have met the BE limits if more subjects had been used[79].

Table 2.1 The Number of Study Subjects Required to Show BE with 80% Power is a Function of WSV and GMR (Sample Size Estimations are for the Case σWT=σWR and σD=0) [79]

Within-subject %CV	GMR (%)	Sample size for a two-way crossover study	Sample size for a four-way fully replicated crossover
15	100	10	6
	105	12	8
	110	20	12
30	100	32	18
	105	38	20
	110	68	36
45	100	66	34
	105	80	42
	110	142	72
60	100	108	56
	105	132	66
	110	236	118
75	100	156	80
	105	190	96
	110	340	172

As MMF is a classic example of drugs with EHC, a large number of models have been developed to characterize the PK profiles of MMF with the emphasis on the modeling of EHC. Figure 2.6 provides a summary of the schemes of previous models with the inclusion of EHC. Because MMF is also a member of BCS class 2 drugs, the variability of systemic appearance cannot be ignored. However, there is no model that focuses on quantifying the variability of MMF systemic

appearance in plasma in the form of MPA to date. Since the systemic appearance PK parameters are good indicators of BE metrics especially Cmax, a model that is designed to explore the WSV of BE metrics (Cmax) from the aspect of a more mechanistic level, aka the intra-individual variability of PK parameters, would aid in developing a novel BE test methodology by better accounting for the WSV.
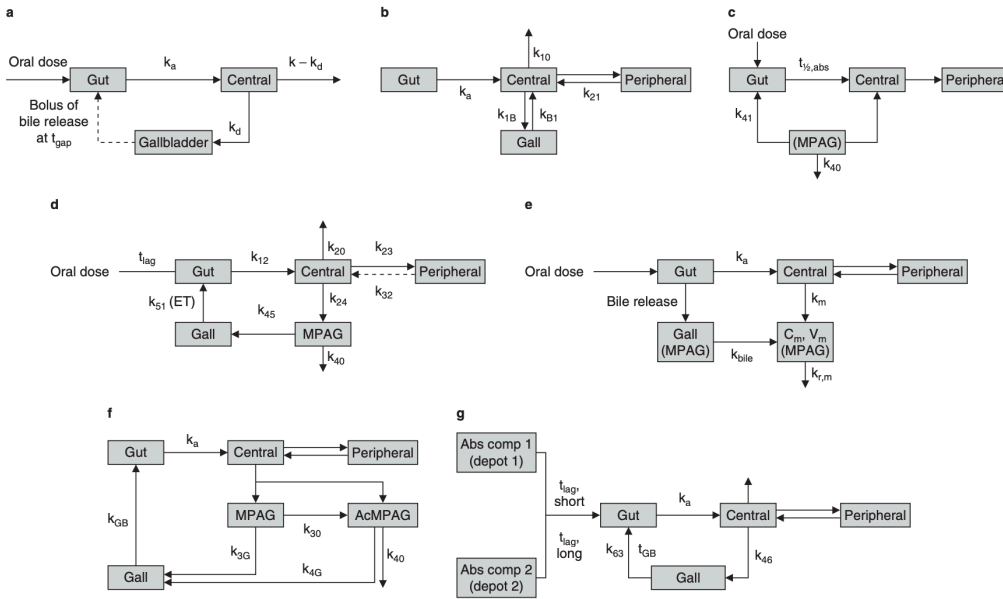
Figure 2.6 Schematics of MMF PK models EHC. 3-compartment EHC model based on a 1-compartment disposition model[82]. (b) Proposed EHC, 2-compartment structural model with first-order absorption with a lag time (tlag)[83]. (c) 4-compartment model, with rate constant describing transfer from fourth to first compartment[84]. (d) Chain compartment model (intestinal, gallbladder, central and peripheral compartments for MPA and central compartment for MPAG[85]. (e) 5-compartment drug and metabolite EHC model with MPA and MPAG plasma concentrations simultaneously[86]. (f) 2-compartment model with linear elimination, with MPAG and MPA acyl-glucuronide (AcMPAG) produced from the central compartment with EHC of MPA via the two metabolites[87]. (g) 2-compartment model accounts for the EHC of MPA. The absorption of MPA was described with two first-order processes with a short and a long tlag and subsequent first-order elimination[88]. Abs comp = absorption compartment; Cm = concentration of MPAG in central compartment; ET = gallbladder emptying time; Gall = amount of MPAG in gallbladder compartment; Gut = amount of MPA in gut compartment; k = first-order rate constant; kxy = transfer rate constant from compartment x to y; ka = absorption rate constant; k_{bile} = biliary excretion rate; k_d = excretion rate constant into gallbladder; k_{GB} = rate constant for the release of recirculated MPA from MPAG and AcMPAG; km = formation rate; $k_{r,m}$ = renal excretion rate of MPAG; $t_{1/2,abs}$ = absorption half-life; t_{gap} = expulsion time of gallbladder; t_{GB} = time of gallbladder compartment opening; V_m = volume of MPAG in central compartment

This analysis is aimed at:

1) characterizing MMF PK profiles especially the absorption phase by developing POPPK models in liver transplanted patients;

2) mechanistically exploring the relationship of the intra-individual variability of MMF PK

parameters with the WSV of MMF BE metrics (Cmax);

3) providing insights into the development of a novel BE test for BCS class 2 HV.

2.2 Method

2.2.1 Clinical study design

A total of 81 adult patients (71 males and 10 females) receiving liver transplantation at Organ

Transplantation Center in Ruijin Hospital were enrolled in the study. The study protocol was

approved by the regional ethics committee (Ruijin Hospital, Shanghai Jiaotong University School

of Medicine) and was performed in accordance with the Declaration of Helsinki and Chinese

guidelines for good clinical practice. All the patients gave informed consent before inclusion. The

demographic and pathophysiological data were collected retrospectively.

The immunosuppressive medications consist of MMF (CellCept®, Roche), tacrolimus (Prograf,

Fujisawa) and corticosteroids. One gram of MMF was given within 6 hours before liver

transplantation. After the surgery, dose of MMF was adjusted based on the occurrence of side

effects. Tacrolimus was orally administrated with an initial dose of 0.1 mg/(kg*day) and was then

adjusted to achieve a steady-state trough concentration of 10-15 mg/ml in the first week and 5-10

ng/ml thereafter. An injection dose of 500 mg of methylprednisolone was given during anhepatic

period and tapered afterwards. After 7 days post first dose, prednisone was used at 20 mg daily

and then further tapered according to the protocol. Blood samples were collected at 0-12 hr after

MMF administration at each visit. The plasma concentrations of MPA and its metabolites were

determined by HPLC assay.

2.2.2 Population PK model

Two population models have been developed based on the data collected from the clinical study above. The first model, which only uses MPA plasma concentration as observations, is intended to focus on the kinetics of systemic appearance, while the second model, which simultaneously model MPA and MPAG plasma concentration-time profiles, was designed for the kinetics of the second peak post-absorption. We assumed there are no impacts from the co-medications on the PK profiles of the MMF. The bioavailability of MMF was assumed to be 1, because the mean absolute bioavailability of oral MMF relative to intravenous MMF based on MPA AUC was reported to be 94% in 12 healthy volunteers[68]. The individual PK parameters at each visit are expressed as below:

$$\theta_{ij} = exp(ln\theta_{pop} + \omega_i + v_{ij}) \qquad (2.1)$$

Where θ_{ij} represents the PK parameter value of the ith individual at visit j;

θ_{pop} represents the population mean of the PK parameter;

ω_i represents the deviation of the individual mean of the PK parameter from its population mean, aka the inter-individual variation (IIV);

v_{ij} represents the deviation of θ_{ij} from the individual mean of the PK parameter, aka the Inter-occasion variation (IOV).

Off diagonal element of the covariance matrix are all zeros.

The expectation maximization (EM) estimation method was employed due to the sparse sampling scheme.

Continuous covariates are scaled by the population median before incorporated as power function in Equation 2.2 and the impact of binary covariate are expressed in Equation 2.3 as below:

$$\theta_i = \theta_{pop}(\frac{COV_i}{COV_{median}})^{\theta_{cov}} \qquad (2.2)$$

$$\theta_i = \theta_{pop} * \theta_{COV}^{COV_i} \qquad (2.3)$$

Where θ_i represents the PK parameter value of the ith individual;

$\qquad \theta_{pop}$ represents the population mean of the PK parameter;

$\qquad COV_i$ represents the base line covariate value of the ith individual;

$\qquad COV_{median}$ represents the population median of the corresponding covariate;

$\qquad \theta_{COV}$ represents the covariate effect.

The metabolism-related kinetics were estimated as parameter(s) based on the data collected from the liver transplanted patients rather than using some pre-determined constants to describe the metabolism process as previous models.

Due to the insufficiency of the data, the impact of covariates was tested by one covariate at a time. The population modeling work was performed on NONMEM in a parallel computing fashion and the visualization plots were generated by R.

2.3 Result

2.3.1 Exploratory data analysis
The demographics and clinical test information of liver transplanted recipients are numerically summarized in Table 2.2 and the corresponding distribution of each item is graphically summarized in Figure 2.7.

Table 2.2 Demographics and clinical test info of liver transplanted recipients

Characteristics	Number or median (range)
Gender (male/female)	71/10
Age (yrs)	49 (14-76)
Weight (kg)	66 (37-87)
ALT	NA
AST	NA
Albumin (g/L)	35 (23-46)
Creatinine (μmol/L)	71 (37-181)
Total bilirubin (μmol/L)	72.7 (21.1-282.1)
HB	NA
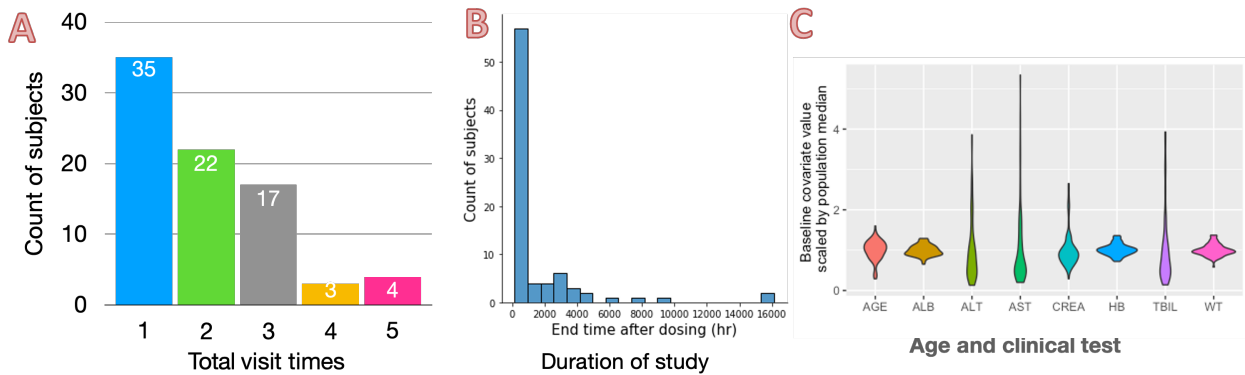
Figure 2.7 Distribution of total visit times (A), duration of study (B), age and clinical test (C). WT: Body weight, ALT: Alanine transaminase, AST: Aspartate transaminase, ALB: Albumin, CREA: Creatine, TBIL: Total bilirubin, HB: Hemoglobin
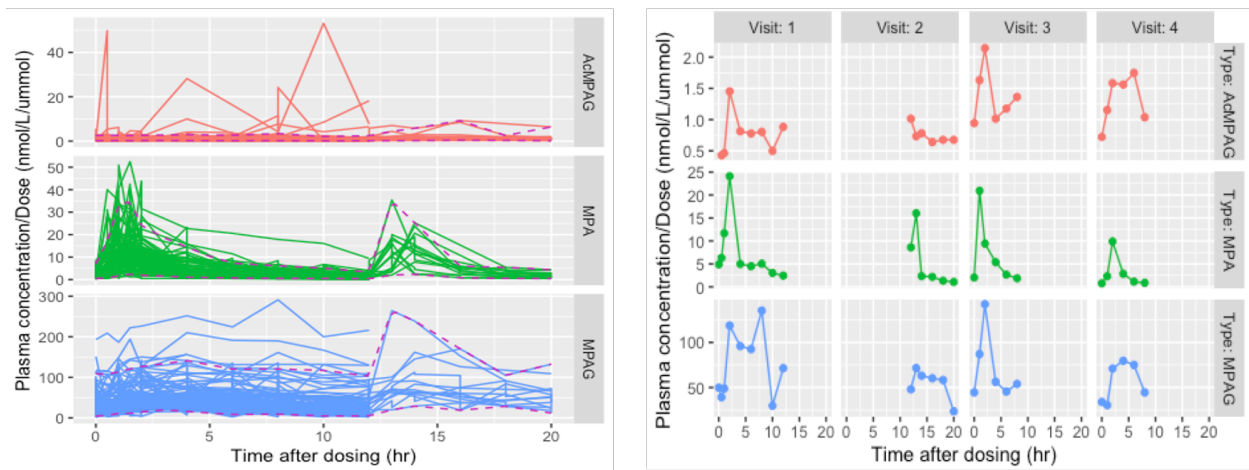
Figure 2.8 Plasma concentration (scaled by dose)-time of MPA, MPAG, AcMPAG of all the subjects at all the visits (left panel) and plasma concentration (Scaled by dose)-time of MPA, MPAG, AcMPAG of Subjects No. 52 at four visits (right panel). The purple red dashed line represents the 95% and 5% percentiles of the scaled plasma concentration at corresponding sampling times.
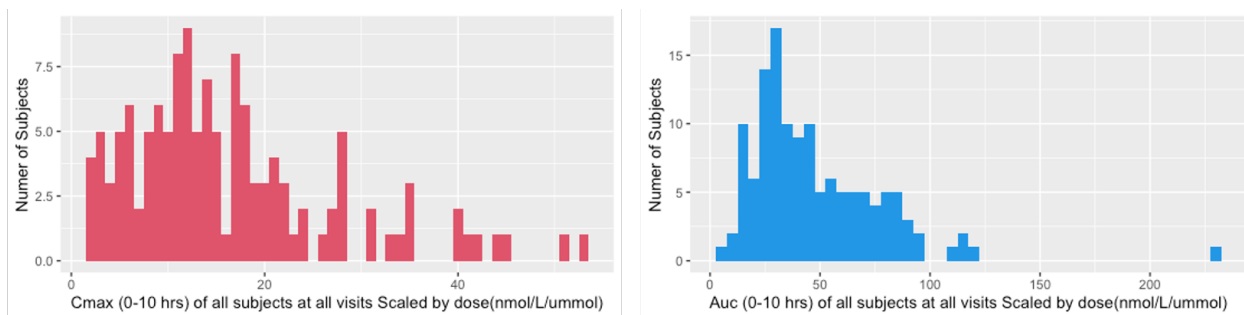
Figure 2.9 Distribution of Cmax0-10 hrs (left panel) and AUC0-10 hrs (right panel) of MPA scaled by dose at all visits for all the subjects.

Table 2.3 The least sample sizes of different study designs if the test and reference MMF have the same CV (0.486), GMR 0.95, 90% CI and the target power is 0.8.

Design	Sample size
2 x 2 crossover study	94
2 x 2 x 4 fully replicated crossover study	48
2 x 2 x 3 partially replicated crossover study	70

A total of 81 liver transplanted recipients were enrolled, among whom 71 were males and 10 were females. 35 of them only had one visits and for subjects who had multiple visits the longest duration of the study could take up to 2 years after the first dose. The distributions of age and clinical test, along with total visit times and duration of study were visualized in Figure 2.7. The left panel of Figure 2.8 showed the overall variability of plasma concentration-time profiles and the right panel displayed the WSV of PK profiles via the example of Subject No. 52. All the plasma concentrations are scaled by the dose due to the different dose for each individual at each visit. A huge difference could be observed when comparing the Cmaxs of MPA for Subject No. 52 at different visits. As shown in Figure 2.9, if only the data at the first 10 hours post dose was extracted, the distribution of Cmax scaled by dose was greatly right skewed with CV 7.03% between subject variability (BSV) and 48.6% WSV, scaled AUC with 6.32% BSV and 40.4% WSV, respectively. Suppose the test MMF generic product has the same WSV as CellCept, GMR is 0.95, CI is 90%

and the target power is 0.8, the least sample sizes of different study designs are calculated in Table 2.3.

2.3.2 Model selection

2.3.2.1 Model 1: just modeling MPA plasma concentration
As MPA is the confirmed pharmacologically active form of MMF, emphasis should be placed on characterizing the kinetics of this compound. Due to the insufficiency of the data, EHC could not be incorporated when just using MPA plasma concentrations as observations. However, since this model was only intended for the systemic appearance (the first peak at the absorption phase) of MMF in the form of MPA, it's acceptable to not include EHC, which is used to describe the second peak. Model fitting was assessed by Akaike information criterion (AIC). As shown in Table 2.4, the best model was NO. 17 model, that is the two-compartment disposition model of MPA with IOV from KA, K20 and V2. Figure 2.10 and Table 2.5 provided the description of the optimal model graphically and numerically, respectively. When checking the effects of baseline covariates on the MMF PK profiles, albumin, total bilirubin, and hemoglobin had significant impact on K20, V2, and KA, respectively, as displayed at Table 2.6.

2.3.2.2 Model 2: simultaneously modeling both MPA and MPAG plasma concentrations
Considering the presence of the second peak, it's still worthwhile to characterize EHC by simultaneously modeling both MPA and MPAG plasma concentrations. The optimal model without the inclusion of IOV in this case was Model No.5, that is MPA two-compartment disposition with lag time and MPAG one-compartment disposition without the gallbladder compartment in EHC. Given the data collected, it's only sufficient to exam the IOV one parameter at a time. As displayed in Table 2.7, inclusion of IOV in terms of Tlag, KA, K23, and K31 could

37

significantly improve the model fitting. Table 2.8 and Figure 2.11 displayed the parameter estimation and the diagram of the optimal model Model No.5, respectively.

Table 2.4 A summary of various models just modeling MPA plasma concentration.

No.	Description	AIC	Note
1	1-COMP	793.250	Residual error 56.4%
2	2-COMP	610.269	K20 IIV 0, Residual error 45.8%
3	2-COMP + Tlag	520.012	Tlag and K20 IIV 0, Residual error 40.5%
4	2-COMP + Tlag IOV	514.894	Tlag and K20 IIV 0, considerable IOV, residual error 37.2%
5	2-COMP + Tlag + KA IOV	499.108	Tlag and K20 IIV 0, considerable IOV, residual error 37.7%
6	2-COMP + Tlag + K20 IOV	427.929	Tlag, K20, and K32 IIV 0, considerable IOV, residual error 34.0%
7	2-COMP + Tlag + V IOV	431.849	K20 and K23 IIV 0, considerable IOV, residual error 34.1%
8	2-COMP + Tlag and KA IOV	492.026	Tlag IIOV 0, negligible K20 IIV, IOV KA > Tlag, residual error 36.7%
9	2-COMP + Tlag and K20 IOV	408.878	Tlag, K20 and K32 IIV 0, IOV K20 > Tlag, residual error 30.3%
10	2-COMP + Tlag and V IOV	419.997	K20, K23, and Tlag IIV 0, considerable IOV, residual error 31.0%
11	2-COMP + Tlag + KA and K20 IOV	395.128	Tlag, K20 and K32 IIV 0, IOV KA > K20, residual error 31.4%
12	2-COMP + Tlag + KA and V IOV	400.812	Tlag, K20 and K32 IIV 0, considerable IOV, residual error 31.1%
13	2-COMP + Tlag + K20 and V IOV	414.239	Tlag, K20 and K32 IIV 0, considerable IOV, residual error 32.9%
14	2-COMP + Tlag, KA and K20 IOV	402.550	Negligible Tlag and K20 IIV, IOV KA > K20 > Tlag, residual error 30.9%
15	2-COMP + Tlag, KA and V IOV	394.630	Tlag, K20 and K23 IIV 0, considerable IOV, residual error 29.9%
16	2-COMP + Tlag, K20 and V IOV	414.686	AIC higher than No. 13
17	**2-COMP + KA, K20 and V IOV**	**392.661**	**Tlag, K20 and K32 IIV 0, considerable IOV, large KA IOV, residual error 30.7%**
18	2-COMP + Tlag + K23 IOV	510.884	Tlag IIV 0, considerable IOV, residual error 38.4%
19	2-COMP + Tlag + K32 IOV	502.127	Small Tlag and K20 IIV, considerable IOV, residual error 35.3%
20	2-COMP + Tlag and K23 IOV	511.032	Negligible Tlag IIV, small Tlag IOV, considerable K23 IOV, residual error 37.6%, AIC higher than No. 18
21	2-COMP + Tlag and K32 IOV	492.238	Tlag IIV 0, negligible K20 IIV, small Tlag IOV, considerable K32 IOV, residual error 34.6%
22	2-COMP + Tlag + KA and K23 IOV	489.308	Tlag IIV 0, considerable KA and K23 IOV, residual error 36.4%
23	2-COMP + Tlag + KA and K32 IOV	480.103	Tlag IIV 0, considerable IOV, residual error 32.5%
24	2-COMP + Tlag + K20 and K23 IOV	429.823	K20 and K32 IIV 0, considerable IOV, residual error 33.5%, AIC higher than No. 6
25	2-COMP + Tlag + K20 and K32 IOV	436.186	K20 and K32 IIV 0, considerable IOV, residual error 34.2%, AIC higher than No. 6
26	2-COMP + Tlag + V and K23 IOV	423.536	Tlag, K20 and K32 IIV 0, considerable IOV, residual error 32.8%
27	2-COMP + Tlag + V and K32 IOV	438.672	K32 IIV 0, negligible K20 IIV, small Tlag IIV, considerable IOV, residual error 34.2%, AIC higher than No. 7
28	2-COMP + Tlag + K23 and K32 IOV	491.730	K23 IIOV 0, small Tlag IIV, considerable IOV, residual error 34.4%
29	2-COMP + Tlag, KA, and K32 IOV	502.635	AIC higher than No. 8
30	2-COMP + KA, V and K23 IOV	397.443	Tlag, K20 and K32 IIV 0, considerable IOV, large KA IOV, residual error 30.4%
31	2-OMP + KA, K23 and K32 IOV	NA	unstable
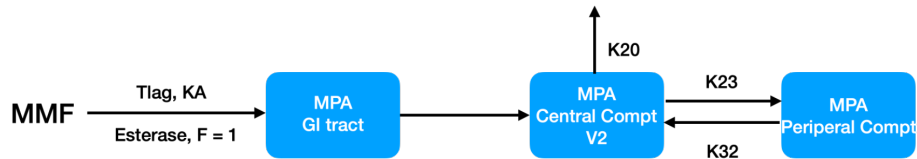
Figure 2.10 Diagram of the best model when just modeling MPA plasma concentrations.

Table 2.5 Parameter estimation of the optimal model just modeling MPA plasma concentration profiles

Parameters	Estimates (95% CI)	IIV (%RSE)	IOV (%RSE)
KA (/hr)	1.10 (0.640-1.89)	0.429 (73.0)	1.18 (76.5)
Tlag (hr)	0.438 (0.415-0.463)	0	0
K20 (/hr)	0.611 (0.403-0.927)	0	0.132 (49.1)
K23 (/hr)	0.481 (0.298-0.775)	0.260 (40.4)	0
K32 (/hr)	0.122 (0.0758-0.198)	0.168 (26.3)	0
V2 (L)	39.25 (25.28-60.95)	0	0.173 (123)
Residual	Estimates (% RSE)		
MPA (%)	30.7 (4.98)		

Table 2.6 A summary of various models checking on the impact of baseline covariates

No.	Description	OBJV
1	ALB-KA	364
2	**ALB-K20**	**363**
3	ALB-V2	392
4	ALB-Tlag	368
5	ALB-K23	391
6	ALB-K32	368
7	SEX-KA	368
8	SEX-K20	370
9	SEX-V2	369
10	SEX-Tlag	376
11	SEX-K23	415
12	SEX-K32	364
13	WT-KA	368
14	WT-K20	369
15	WT-V2	367
16	WT-Tlag	379
17	WT-K23	368
18	WT-K32	368
19	CREA-KA	367
20	CREA-K20	366
21	CREA-V2	367
22	CREA-Tlag	521
23	CREA-K23	371
24	CREA-K32	369
25	ALT-KA	373
26	ALT-K20	400
27	ALT-V2	373
28	ALT-Tlag	369
29	ALT-K23	373
30	ALT-K32	373
31	AST-KA	371
32	AST-K20	383

33	AST-V2	372
34	AST-Tlag	373
35	AST-K23	374
36	AST-K32	373
37	TBIL-KA	365
38	TBIL-K20	368
39	**TBIL-V2**	**363**
40	TBIL-Tlag	372
41	TBIL-K23	365
42	TBIL-K32	398
43	**HB-KA**	**363**
44	HB-K20	380
45	HB-V2	368
46	HB-Tlag	373
47	HB-K23	374
48	HB-K32	369

Table 2.7 A summary of various models simultaneously modeling both MPA and MPAG plasma concentrations

No.	Description	OBJV	Note
1	2-COMP MPA + 1-COMP MPAG	306	Negligible K20 IIV, residual: MPA= 49.4%, MPAG = 19.5%
2	2-COMP MPA + 1-COMP MPAG + EHC + No Gallbladder	300	Negligible K20 IIV, K31 IIV = 0, VMPAG IIV = 0, residual: MPA= 49.3%, MPAG = 19.5%
3	2-COMP MPA + 1-COMP MPAG + EHC + Gallbladder	299	Negligible K20 K34, and K41 IIV, PMPAG = 97.7%, VMPAG IIV = 0, residual: MPA= 49.1%, MPAG = 19.5%
4	2-COMP MPA + 1-COMP MPAG + Tlag	245	Negligible K20 IIV, Tlag IIV = 0, residual: MPA= 44.4%, MPAG = 20.2%
5	2-COMP MPA + 1-COMP MPAG + Tlag + EHC + No Gallbladder	226	Negligible K20 IIV, K31 and VMPA IIV are 0, PMPAG = 97.6%, residual: MPA= 43.0%, MPAG = 19.6%
6	2-COMP MPA + 1-COMP MPAG + Tlag + EHC + Gallbladder	230	Negligible K20 K34, and K41 IIV, PMPAG = 99.2%, VMPA and VMPAG IIV = 0, residual: MPA= 43.2%, MPAG = 19.6%
7	2-COMP MPA + 1-COMP MPAG + Tlag + EHC + Tlag with IOV	212	Negligible K20, K23, and K31 IIV, Tlag: IIV = 0 and considerable IOV, PMPAG = 97.3%, residual: MPA= 40.6%, MPAG = 19.6%
8	2-COMP MPA + 1-COMP MPAG + Tlag + EHC + KA with IOV	187	Negligible K20, K31, and K30 IIV, KA: IOV significantly higher than IIV, PMPAG = 96.9%, residual: MPA= 39.8%, MPAG = 19.2%
9	2-COMP MPA + 1-COMP MPAG + Tlag + EHC + K23 with IOV	66.0	Negligible K20 and K31 IIV, low Tlag IIV, K23 IOV significantly higher than IIV, PMPAG = 99.4%, residual: MPA= 37.1%, MPAG = 16.4%
10	2-COMP MPA + 1-COMP MPAG + Tlag + EHC + K20 with IOV	62.8	Negligible K31 and Tlag IIV, K23 IIV = 0, K20 IOV significantly higher than IIV, abnormally low PMPAG = 3.97%, residual: MPA= 43.4%, MPAG = 16.0%
11	2-COMP MPA + 1-COMP MPAG + Tlag + EHC + K31 with IOV	-1.52	Negligible K20 IIV, PMPAG = 97.8%, K23 IIV = 0, K31 low IIV and considerable IOV, residual: MPA= 40.2%, MPAG = 13.7%
12	2-COMP MPA + 1-COMP MPAG + Tlag + EHC + K30 with IOV	65.5	Negligible K31 and VMAPG IIV, low Tlag IIV, abnormally low PMPAG = 16.2%, K20 IOV significantly higher than IIV, residual MPA = 43.4%, MPAG = 16.0%

PMPAG = K23/(K23 + K20)

Table 2.8 Parameter estimation of the optimal model simultaneously modeling both MPA and MPAG plasma concentrations

Parameters	Estimates (95% CI)	IIV (%RSE)
KA (/hr)	0.700 (0.562-0.871)	0.315 (38.1)
Tlag (hr)	0.380 (0.310-0.466)	0.324 (22.8)
K20 (/hr)	0.0284 (0.0265-0.0305)	0.0000291
K23 (/hr)	1.14 (0.937-1.38)	0.221 (21.5)
K25 (/hr)	1.03 (0.621-1.72)	1.32 (33.6)
K52 (/hr)	0.0120 (0.00681-0.0213)	1.89 (38.6)
K31 (/hr)	0.0474 (0.0300-0.0758)	0
K30 (/hr)	0.125 (0.0973-0.160)	0.324 (22.8)
V2 (L)	21.1 (17.5-25.5)	0
V3 (L)	11.0 (8.67-14.0)	0.135 (52.0)
Residual	Estimates (% RSE)	
MPA (%)	43.0 (4.84)	
MPAG (%)	19.6 (4.90)	
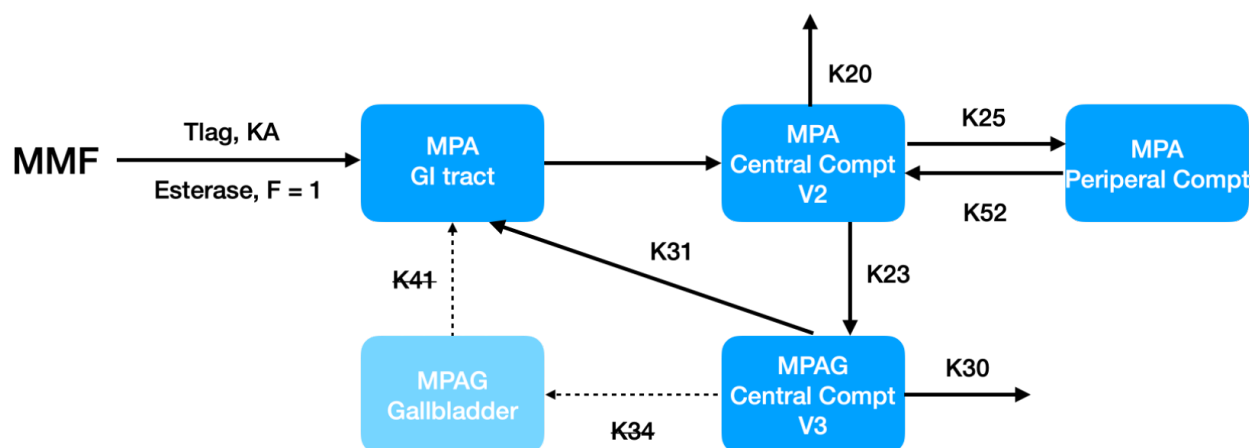
Figure 2.11 Diagram of the models when simultaneously modeling MPA and MPAG plasma concentrations. The flow of the optimal model Model No.5 was represented with solid arrows. The light blue square and dashed arrows are the gallbladder compartment and the corresponding kinetics used in Model No.6.

2.3.3 Model evaluation

The diagnostic plots of Model 1 look not so perfect mainly due to exclusion of the EHC (Figure 2.12). A large impact of the random effects can also be anticipated.
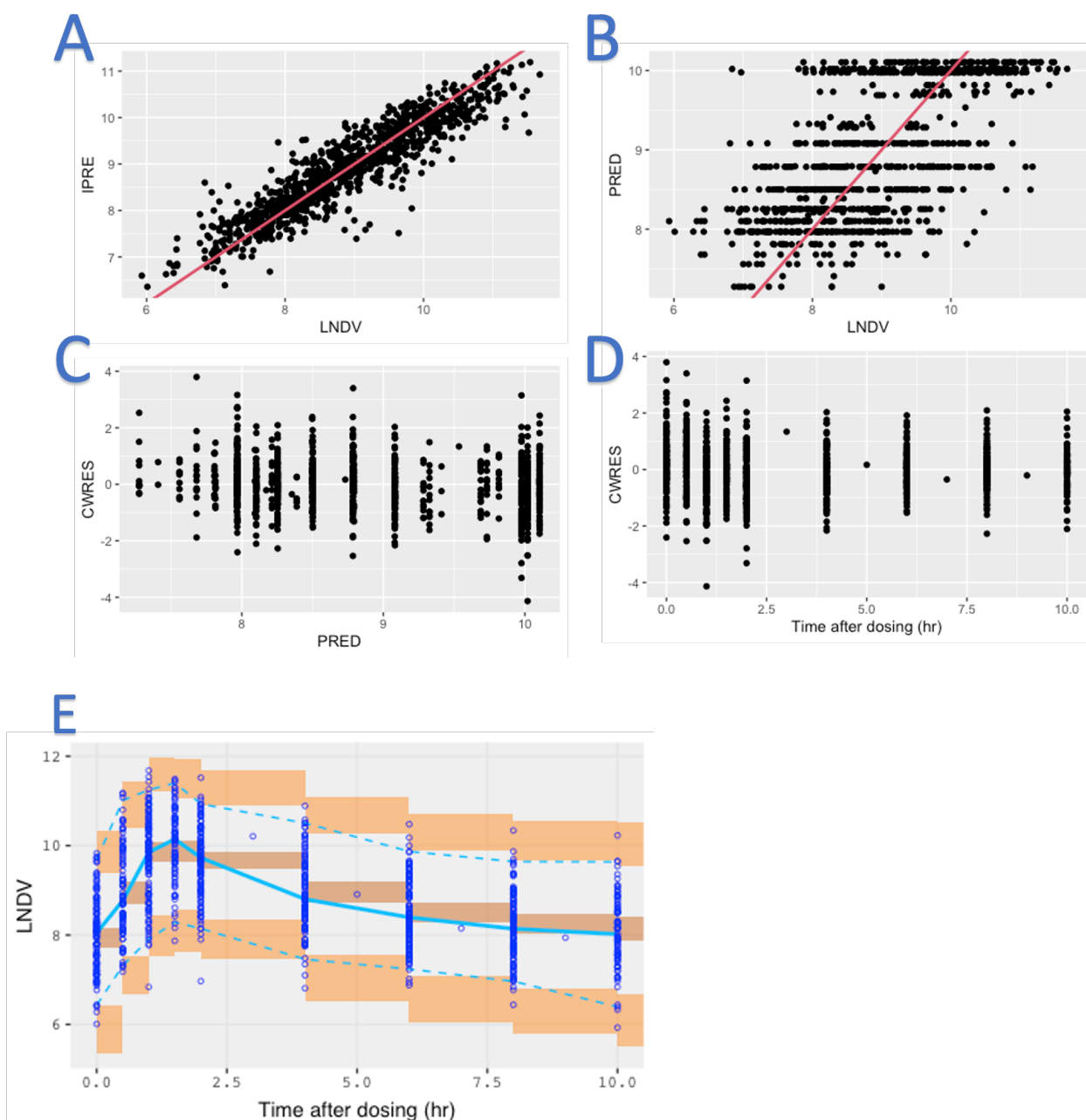
Figure 2.12 Goodness-of-fit plots (A-D) and visual predictive check (VPC) plot (E) when just modeling MPA without EHC. LNDV: logarithm of the observed MPA plasma concentration; IPRED: individual prediction of the MPA plasma concentration after taking the logarithm. PRED: logarithm of the population prediction of MPA plasma concentration. CWRES: conditional weighted residual errors. The solid blue line represents the median observed MPA plasma concentration, and the dark orange area around it represents a simulation-based 95% confidence interval for the median. The observed 5% and 95% percentiles are presented with dashed blue lines, and the 95% confidence intervals for the corresponding model predicted percentiles are shown as the corresponding blue fields.

2.4 Conclusion and discussion

Considerably more attention should be drawn to how to better conduct BE studies for HV drugs. Firstly, HV drug products constitute a disproportionately high percentage of drug product submissions among the failed BE studies in one of the FDA statistics[79]. While HV drugs only appear to represent about 20% of ANDAs evaluated by FDA, 45% of the failed BE studies reviewed at FDA from 2009 to 2012 belong to HV drugs based on Table 2.9[79]. Secondly, the current approaches to addressing the high failure rate issue of HV drug are economically and scientifically unfriendly. As summarized in Table 2.10, among the 45% of the failed BE studies mentioned above, 91% of the initially failed ANDAs met the BE limits by increasing sample size in two-way crossover BE study and 7% of them passed the BE tests by changing to three-way study design and RSABE[79]. The large sample size required would increase the costs of generic drugs, perhaps more significantly lead to unnecessary human testing, and deter the development of new generic products. Furthermore, the scientific debate of setting the regulatory constant and cutoff for RSABE limits the wide application of this approach in BE studies[79]. Therefore, a novel methodology that is targeted to solve the high WSV of BE metrics (Cmax, AUC) is highly desirable.

Table 2.9 Number of Failed BE Studies Reviewed at FDA's OGD Since the "All Bioequivalence Studies" Rule Became Effective in July 2009 till 2012. [79]

Description	BE Studies		ANDAs	
	Number	Percent of total	Number	Percent of total
WSV of AUC and $C_{max} \geq 30\%$	92	45	45	37
WSV of AUC and $C_{max} < 30\%$	113	55	76	63
Totals	205	100	121	100

Table 2.10 Changes Made to the Study Design or Formulation to Achieve Successful Pivotal BE Studies of HV Drugs after the Initial BE Studies Failed to Meet the Acceptance Criteria[79]

Change made	No. of ANDAs	Percent
Increase in sample size two-way crossover BE study	41	91
Changed to three-way study design and RSABE	3	7
Reformulated	1	2
Total	45	100

The WSV of BE metrics (Cmax and AUC) can be largely explained by the intra-individual variability of systemic appearance related PK parameters in the case of HV drugs. As we know, the rate and extent of drug dissolution and subsequent absorption rely on the characteristics of GI physiological environment such as hydrodynamics, gastric emptying rate, intestinal transit, GI pH, buffer capacity and the fasted vs fed state of the GI tract, as well as the physicochemical properties of the drug and the properties of dosage forms[89,90]. Hence, a slight change in physiological factors may give rise to significant variabilities in the process of drug oral uptake and systemic appearance. This may be especially pronounced for drug molecules suffering from a low aqueous solubility but showing a high intestinal permeability, classified as BCS class 2 compounds[5], which are likely to be HV drugs. For the IR products of these HV drugs, dissolution could be the rate-limiting step of their absorption, rendering them super sensitive to environmental changes at absorptive sites. Consequently, high variabilities are commonly observed in their dissolution and absorption profiles even for the same individual at different measuring times (high intra-individual variability) and contribute greatly to the WSV of BE metrics. These high variabilities can be represented as the intra-individual variability of systemic appearance from the aspect of PK. Simulation of luminal dissolved drug concentration based on GI physiology is therefore one of the key determinants of reasonable prediction of drug absorption.

The intra-individual variability of systemic appearance related absorption PK parameters can be partly attributed to different states under fasted conditions. The dynamics of GI transit and luminal

pH fluctuation and the underlying physiological basis has been described in detail in Section 1.3 and 1.2.2, respectively. Oral dosing under the fasted conditions can occur at one of the very different states of MMC, and thus results in high variability.

While under postprandial conditions, the motility and secretion pattern vary greatly from those under inter-digestive conditions. The stomach mechanically breaks down the ingested food into small particles to enhance digestion like a homogenizer. The pylorus remains closed for prolonged periods and acts as a sphincter that controls the amount and size of food particles that can exit the reservoir, the stomach, in the intermittent partial openings. The gastric emptying rate depends on both the physical and chemical characteristics of the meal with a rate of 200 kcal/hr on average: liquids empty most rapidly; solids empty only after a lag phase. The cutoff particle size to be held back or to be delivered to the duodenum is 1-2 mm. Meanwhile, motility events in the small intestine become more frequent, with patterns designed to mix the meal with intestinal secretions and to maximize exposure of the digested nutrients to the absorptive mucosa. Intestinal transit is slowed in proportion to the number of calories presented to the intestine. The gallbladder serves to store and concentrate bile coming from the liver in the period between meals. Postprandial gallbladder contraction coincides with gastric emptying. The entry of the meal into the duodenum triggers the secretion of a bolus of concentrated bile with the entry of dietary lipids into the small intestine[10].

In this study, two population PK models have been developed with the emphasis on exploring the variability of systemic appearance and EHC of MMF. The first model only employed MPA plasma concentrations as observation and investigated the intra-individual variability of PK parameters in the form of IOV. The second model used both MPA and MPAG plasma concentration profiles to

further check EHC. The effect of IOV of PK parameters was checked only one at a time in the second model due to data insufficiency. The results of both two models showcased that the intra-individual variability plays a predominant role instead of IIV in terms of variability of absorption-related PK parameters. The high intra-individual variability of absorption-related PK parameters would be expected because the timing of drug administration and sampling is a mix of fasted and fed states and there is no knowledge of motility state and MMC phases at the time of dosing and the variable duration of MMC cycle.

Surprising as it may look at first glance, GE is likely to be the rate-limiting step in the systemic appearance process of CellCept and hence accounts for the intra-individual variability of its systemic appearance. MMF has to go through the disintegration, release and dissolution in the stomach, gastric emptying, transport into the mucosal cell, and conversion to MPA before showing up in the circulating system in the form of MPA. Given the fact that 1) CellCept is the only drug product of MMF used in this study; 2) CellCept is IR by design; and 3) rapid and stable dissolution in the stomach reported by Scheubel et al[91], negligible variation should be anticipated from the release and dissolution process of CellCept. Also considering the high permeability and super low pKa value of the morpholino group, it would be natural to speculate that CellCept is quickly transported into the mucosal cells of the proximal small intestine lumen, once being dissolved and emptied from the stomach. Finally, due to a complete and rapid pre-systemic metabolism, the conversion of MMF to MPA would not introduce significant further variation. Taken together, all the systemic availability events of CellCept are controlled by gastric emptying and Cmax can be seen as a function of gastric emptying. The indirect evidence in a previous publication showed that plasma Cmax/AUC was negatively associated with time to MMC phase 3 post dosing in the fasted state after a dose of Ibuprofen[92]. Therefore, it seems mechanistically correct to claim that intra-

individual variability of MMF systemic appearance is essentially the intra-individual variability of GE.

The dominance intra-individual variability possessed over IIV in terms of absorption-related PK parameters would suggest the BE studies of MMF from two different perspectives. Firstly, dosing under fed conditions is recommended for the *in vivo* clinical BE studies. Because less WSV of BE metrics is anticipated under fed conditions with the same meals distributed to each subject compared with under fasted conditions due to the randomness of dosing and sampling time relative to the three MMC phases, assuming there are no food-drug interactions. In addition, the bile secretion rate is well regulated under fed conditions to reduce the variations in EHC. Secondly, a novel predictive *in vitro* BE test only measuring product difference would be a new trend because *in vivo* clinical BE studies are greatly confounded by the large GI variations. Obviously, the impact of BSV is removed from an *in vitro* test to simplify the analysis. In the case of MMF, as discussed previously, the WSV of systemic appearance, though large, is actually the intra-individual variability of GE, and the product difference is significantly confounded with random GI motility. Despite of the huge intra-individual variability of EHC, it only explains the variation of the second peak in post absorption phase, which is also related with the GI physiological dynamics and the time of food consumption post dosing. Thus, a predictive dissolution test comparing the profile of the test formulation to that of the innovator MMF product with sufficient simulation of Gastric emptying dynamics would be a desirable surrogate. If the release and dissolution rate of the test MMF formulation is rapid like the original CellCept (85% dissolved within 15 mins), then equivalent dissolution profiles would indicate bioequivalence of the test and reference MMF. As for other HV drug products, an *in vitro* dissolution test sufficiently simulating GI physiological dynamics for oral systemic availability (absorption) would also serve the BE test purpose.

While the predictive *in vitro* BE test has gained increasing attention since the development of the BCS classification system, it is not widely accepted even today. This is because current dissolution tests do not take full advantage of the extensive available GI physiologic information. For instance, the static fasted intestinal pH values of FaSSIF (6.5) and USPSIF (6.8) are embodied as a strong buffer capacity, that is a 5 and 7.7 times higher capacity than human intestinal fluid, respectively[93,94], which apparently fails to reflect the real *in vivo* GI physiological conditions, let alone to capture the characteristics of drug dissolution, absorption and systemic exposure kinetics (Cmax, AUC, Tmax). These routine methodologies are more suitable for quality control purposes and are only used as tools to understand the effects of formulation and processing changes[95], so a new methodology incorporating the physiologically relevant GI information is desired. A good starting point would be a typical fasted BE condition, but could evolve into tests simulating the GI environment of patients and special populations. Some *in vivo* tests may be required for final assurance of performance in patients, but the *in vitro* dissolution test would be the mainstay of product development and quality control.

Although there's a sea of MPA population models attributed to EHC of MPA[96], the uniqueness of the two models developed in this study lies:

1) for the first time focusing on the complicated absorption process of MMF as a BCS class drug, relating the intra-individual variability of the systemic appearance PK parameters to the WSV of BE metrics and mechanistically explaining the intra-individual variability from the aspect of GI physiology dynamics;

2) introducing IOV to the metabolism process rather assume a static MPA clearance rate with

fixed ratio of converting to MPAG: being cleared via other routes, considering the subject are liver transplanted recipients and time-varying Cmax and AUC have been noted among these type of patients in FDA drug label of CellCept[68].

Significant conclusion can be drawn with regard to the complicated MMF absorption process by leveraging the data at hand, however, the variance in the population model could be better quantified if:

1) more frequent sampling schema was designed for the absorption phase;

2) food intake time relative to MMF dosing was well recorded;

3) the clinical biochemistry covariates were collected at least at each visit;

4) the protein-bounded MPA were determined rather than just the total MPA, since MPA is highly bounded to albumin.

In the future, we would like to dive deep into the mechanisms of HV drug absorption process by further exploring the relationship between GI physiology dynamics and PK variability. Possible attempts include utilizing:

1) SmartPill to record local GI pH fluctuation and motility;

2) tubes in the GI tract to directly determine the concentrations of dissolved and undissolved drug;

3) a comparison of the radiolabeled and cold-treated drug.

2.5 Acknowledgement

Chapter 3 Text Mining for Drug BCS Classification Using Natural Language Processing

3.1 Introduction

BCS theory employs the physicochemical concept of maximal solubility in GI physiological pH range relative to the highest strength and intestinal permeability to represent the behavior of an IR solid drug dosage form in the GI tract and further predict the systemic appearance profiles of this oral drug product. It has been widely adopted as the criteria for the biowaiver of a generic IR solid drug product. FDA has launched a web-based application to facilitate collecting information of FDA-approved drug products without the inclusion of drug BCS class. (https://nctr-crs.fda.gov/fdalabel/ui/search) Considering the enormously high volume of documentations regarding drugs designed for oral absorption, it is desirable to harvest BCS classification information from various sources online to support pharmaceutical development and clinical decision.

However, the information extraction (IE) of drug BCS class is never an easy task. Firstly, the relevant solubility profiles and intestinal permeability are scattered in various documents like drug labels, regulatory review documents, and literatures. Secondly, a large number of the materials are uploaded online as a scanned pdf file, which means people cannot find the BCS info by just googling. Thirdly, compared to the structured data, drug BCS class and relevant info are usually documented in a more natural, expressive and unstructured way. Fourthly, biomedical documents

usually involve heavy use of domain-specific terminology and knowledge, making it difficult for a non-pharmaceutical scientist to handle it directly.

Natural language processing (NLP), a confluence of computer science and linguistics, enables the extraction of BCS class from a deluge of information and the transformation of the text into actionable data that can be quickly visualized and analyzed at different stage of drug development via ML/AI algorithm. Some of the major areas of NLP include but not limit to question answering system, automatic summarization, machine translation, speech recognition, and document classification[97]. Recently, it also gains popularity among big pharmaceutical companies and bio-tech organizations. Just to name a few, NLP has been used for effective search for patent landscaping and competitive intelligence in Pfizer, discovery potential novel biomarkers and phenotypes for diabetes and obesity, from PubMed, clinical trial data, and internal research documents in Merck, and text mining EMRs for patient stratification of heart failure risk in BMS[98].

The goal of this project is to propose a framework to use text mining technology mainly NLP to: extract of the solubility, permeability and/or BCS class information of IR solid oral drug products from drug labels, FDA review documents, FDA orange book, and some selected biomedical literatures;

1) identify drug BCS class based on the text mining results in a supervised way;
2) help build a database carrying drug oral absorption information to further fulfill the BCS potential in the development of oral drug products.

3.2 Methods

3.2.1 Documents retrieval

The scope of the BCS source search included appropriate drug product lists with its active pharmaceutical ingredients (APIs), highest strength, text corpora consisting of drug labels, FDA review documents, and literatures, knowledge of pharmaceutical sciences. First, search the FDA orange book (published on 03/23/2020) thoroughly to get a list of product name, active pharmaceutical ingredient (API) and the corresponding highest strength with the inclusion criteria regarding the drug products: 1) still on the market; 2) with known strength; 3) oral dosage form; 4) immediate release; 5) solid. Secondly, query the FDALabel database based on the product names acquired in the first step and save the query results; Thirdly, filter the query results from the second step with same inclusion criteria mentioned in the first step to obtain all the application numbers of the drug products; Fourthly, Crawl the Drugs@FDA website (https://www.accessdata.fda.gov/scripts/cder/daf/) automatically according to the application numbers compiled in the third step and save all the pdfs files mainly all the labels and FDA review documents. Fifthly, add the necessary biomedical literatures found on Web of Science, Pubmed, and Research Gate to the large corpus in the fourth step to form a collection of documents, if there are no solubility, permeability or BCS class information on the FDA website. Sixthly, narrow down the scope of drug products by manually inspecting 1) it's not absorbed in the oral cavity; 2) BCS class has been confirmed by FDA reviewer; 3) it has clear supporting solubility and permeability data. Seventhly, convert all the scanned pdf files to plain text for the text data and to Microsoft Word for the tabular data using the Optical character recognition (OCR) feature provided by Adobe Acrobat to get the raw large corpus for NLP analysis. Eighthly, label the BCS

class of each drug substance manually based on my knowledge of pharmaceutical sciences and save the classification result as a gold standard for the ML/AL classification models.

3.2.2 Text preprocessing

Sentence splitting, tokenization and stemming was done sequentially to the raw text corpus. Only the sentences satisfying the following conditions were kept:

- Containing "BCS", "Biopharmaceutics Classification System",

- Containing "solub" and "pH"

- Containing "permea" and "%"

- Containing "bioavailability" and "%" but no "relative", containing "bioavailability" and "fraction" but no "relative", containing "absor" and "gastrointestinal", containing "absor" and "GI", containing "absor" and "complete"

- No "rat", "mice", "dog", "monkey","pig", etc

Assign 80% of the drug substances and the corresponding tokenized sentences to the training dataset.

3.2.3 ML/AI model

The pre-processed text was then transformed into quantifiable data by the algorithm "word2vec". ML algorithms Logistic regression (LR), random forest (RF) and Bi-LSTM were employed to train the model and the best models achieved from the training results were applied to the test set to predict BCS class.

3.2.4 Ontology construction

The BCS ontology is composed of three components: solubility, permeability, and BCS class.

"*Solubility*" specified the solubility profiles numerically (eg. 2.5 mg/mL) or textually (e.g insoluble) under certain pH ranges (eg. pH range 1-6.8, physiological pH range, pH 3). "Permeability" indicated drug permeability result (eg. high permeability, 100%) in the corresponding form (eg. Caco-2 cells, absolute bioavailability). "*BCS*" represents the direct claim of drug BCS class found in the text like BCS Class 2, Class III of the Biopharmaceutical Classification System, et.al.

3.2.5 IE using NER

Ontology-based name entity recognition (NER) enable to better locate the BCS-related named entities to greatly narrow down the scope for analysis.

The drug BCS class information is extracted as tagging entities as below:

- Drug tagging: a drug substance library was built based on the list of API and tagged as like <DRUG: mycophenolate mofetil>

- Highest strength tagging: the highest strength of a drug product was acquired from the list described in 3.2.1 in the form <DOSE: 500 mg>

- Solubility tagging: <SOL_RE: 4.27 mg/mL> at <SOL_PH: pH 3.6>

- Permeability tagging: In 12 healthy volunteers, the mean <PER_SU: absolute bioavailability> of oral mycophenolate mofetil relative to intravenous mycophenolate mofetil (based on MPA AUC) was <PER_RE: 94%>.

- BCS class tagging: MMF is a weak base classified as a <BCS: BCS Class II> substance.

3.2.6 Identification of BCS class with NER

The priority of determining BCS class for a drug product is listed as follows:

1) If a BCS info is tagged directly, then read the number within the BCS class tagging;

2) If there's no BCS class tagging for a drug substance, then check the solubility:

- If any numerical results was tagged as SOL_RE then compared the results with highest strength (mg) divided by 250 mL to get a binary result in term of solubility either "high" or "low";

- If the content within the tag SOL_RE doesn't contain any numbers, then process the text via NLP ML/AI algorithm;

3) If there's no BCS class tagging for a drug substance, then check the permeability.

- If any numerical results tagged as PER_RE were higher than 85%, then "high" permeability; otherwise "low";

- If the content within the tag PER_RE doesn't contain any numbers, then process the text via NLP ML/AI algorithm;

4) If there's no BCS class tagging for a drug substance, then identify the BCS class for a drug substance based on the results from step 2 and step 3.

3.2.7 Evaluation

The prediction results were evaluated with gold standard described in the eighth step of 3.2.1 with the metrics precision, recall, and F1 scores.

3.3 Results

3.3.1 Exploratory data analysis

There are 38118 drug products in the FDA orange book published on 03/23/2020. 21114 of them are still on the market. Among these, 21057 drug products with known strength. 14641 of them are designed as oral dosage form and 12186 are IR. As shown in Figure 3.1, 820 of the distinct

APIs collected from the 11232 solid dosage forms for further analysis. Based on the BCS class, class 2 and 3 both occupy around 30% of the APIs. Class 1 is around 25% and class 4 is less than 20%.
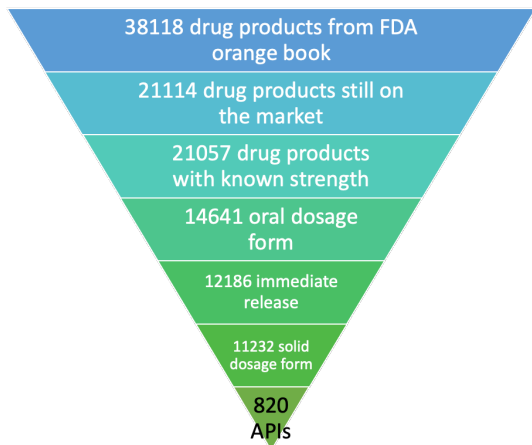

Figure 3.1 Summary of drug products in FDA orange book

Figure 3.2 displays how the BCS ontology is tagged in a text corpus after pre-processing.
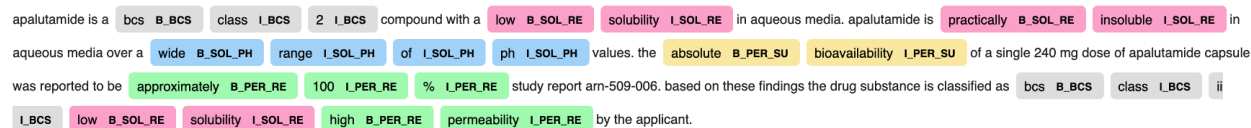
Figure 3.2 An example of annotated corpus based on BCS ontology.

3.3.2 Model Evaluation

Table 3.1 showcased that NER using the BCS ontology performed significantly better predictions compared with handling the text just with ML/AI models.

Table 3.1 Evaluation results of the ML/AI and NER models

Method	Class 1			Class 2			Class 3			Class 4		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
LR	0	0	0	0.33	1	0.5	0	0	0	0	0	0
RF	0	0	0	1	1	1	0.5	1	0.67	1	1	1
Bi-LSTM	0	0	0	0	0	0	0.25	1	0.4	0	0	0
NER	1	1	1	1	1	1	1	1	1	1	1	1

3.4 Conclusion and discussion

NER-based NLP did a nice job on extracting information and identifying drug BCS class from unstructured data on a large scale. Classification based on NER is better than classical machine learning algorithm combined with deep neural network models.

A well-annotated BCS corpus and ontology can facilitate text mining in the development and practice of drug substances in IR solid oral dosage forms on a large scale. In the future, we are interested in expanding the NLP approach to the whole drug oral absorption arena like different formulations, prodrugs, food effects, absorption models.

Appendix

The appendix contains a summary of tangential projects pursued during my PhD studies, though not fully completed due to various reasons. They are summarized below for future reference.

Project 1: Stochastic Differential Equation-Based Hierarchical Model of the Fluid Volume in the Stomach

Introduction

The rate and extent of solid oral drug products release, dissolution and absorption depend largely on the dynamic fluid volumes along the gastrointestinal (GI) tract. Since the fluid volume in the stomach is determined by saliva and gastric secretion, gastric emptying and oral dosing, a model built based on the observed fluid volumes in the stomach at given sampling times can be used to describe the kinetics of these physiological and drug product related process. This is the basis for a more mechanistic approach to predicting drug release, dissolution and subsequent absorption profiles. The predictive results are particularly important for the monitoring of highly variable drugs with narrow therapeutic windows and the design of bioequivalence studies. Conventionally, these physiological parameters are all treated as time-invariant and the fluid volume in the stomach is not considered explicitly, however, recent research results demonstrated these physiological processes are very dynamic[26-31].

Stochastic differential equation (SDE) -based hierarchical model is preferred in the dynamic system to provide a robust parameter estimation, and serves as a diagnostic tool for structural model misspecification or parameter fluctuations. In contrast to the classical ordinary differential equation (ODE) based population approach, where intra-individual variability is only assigned to measurement errors, the employment of SDE method allows for decomposition of the intra-individual variability into a system noise term arising from time-dependent or serial correlated errors like unknown or incorrectly specified dynamics and a measurement noise term accounting for uncorrelated errors such as assay error[99].

The aims of this project are (1) to capture the great inter- and intra- individual variability of the stomach fluid volume and transit in healthy subjects; (2) to find out the physiological parameter(s) representing the dynamics in the stomach by implementation of SDE method and comparing the results with ODE.

Methods

Study design

Twelve healthy subjects were enrolled and had their gastric and small intestinal water volume measured by magnetic resonance imaging (MRI) scans before and after consumption of 240 mL of water at pre-determined intervals for up to 2 hours post-ingestion as shown in Appendix Figure 1. The SDE- and ODE- based mixed effects morels were implemented and compared using Extended Kalman Filter (EKF) algorithm with a 3*3 prediction variance-covariance matrix in NONMEM.
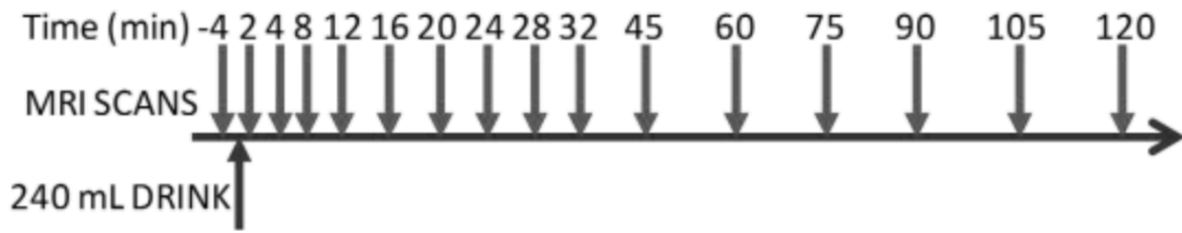
Appendix Figure 1: Study diagram

Model assumption

The stomach compartment was basically modeled with zero-order saliva and gastric secretion coefficient Kse and first-order gastric emptying coefficient Kge as depicted in the Appendix Figure 2.
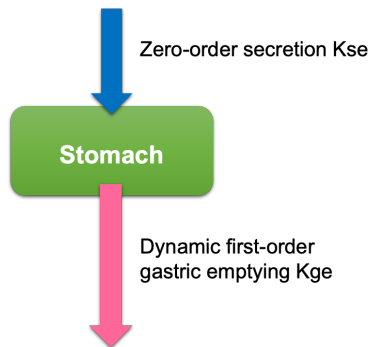
Appendix Figure 2: Schematic illustration of the structure model for the fluid volume in the stomach.

The first-stage model with SDEs written in state-space form consists of a set of continuous time system equations in Equation (1)-(2), and discrete time measurement equations in Equation (3), that is,

For subject i at jth sampling time:

Continuous time state space stochastic differential equations:

$$dXstomach_i(t) = (-Kge_i(t) \times Xstomach_i(t) + Kse_i)dt \hspace{2cm} (1)$$

$$dlogKge_i(t) = -\alpha(logKge_i(t) - logKge_i*) + \sigma_{wKge}dW_i(t),\ W_i(t) - W_i(s) \sim N(0, |t\text{-}s|I) \quad (2)$$

Discrete time measurement equations:

$$Vstomach_{ij} = Xstomach_{ij} \times (1 + \sigma_{prop}\varepsilon1) + \sigma_{add}\varepsilon2,\ \varepsilon1\ and\ \varepsilon2 \sim N(0,1) \quad\quad\quad (3)$$

Inter-individual variability:

$$Kge_i* = Kge* \times exp(\eta_{Kgei}),\ \eta_{Kgei} \sim N(0, \omega_{Kge}^2) \quad\quad\quad\quad\quad (4)$$

$$Kse_i* = Kse* \times exp(\eta_{Ksei}),\ \eta_{Ksei} \sim N(0, \omega_{Kse}^2) \quad\quad\quad\quad\quad (5)$$

where Xstomach is the stomach fluid volume state variable, Vstomach is the observation variable, t is the time variable, $\sigma_w dW$ is the system noise, I is the identity matrix, $\sigma_{prop}\varepsilon1$ and $\sigma_{add}\varepsilon2$ are the heteroscedastic and homoscedastic measurement errors, respectively.

σ_w is a scaling diffusion term, where W is a standard Wiener process. The possible fluctuations of parameter were checked in the scenario of dynamic (1) Kse σ_{wKse}, (2) Kge σ_{wKge}, (3) Kse and Kge σ_{wKse} and σ_{wKge} by Ornstein-Uhlenbeck diffusion process. We just showed the case of time-varying Kge as an example below in Equation (2). If the diffusion term σ_w is zero, the SDE reduces to an ODE. The usual physiological interpretation of the parameters is thereby preserved in the SDE model formulation.

The second-stage model describing the inter-individual variability (IIV) is included in the same way as for ODEs. The typical individual parameter value are modeled in Equation (4) and (5) as a function of the fixed-effect population parameters and random-effect parameters η. The random-effects η are assumed independent and normal distributed with zero mean and standard deviation

ω. The three levels of random-effects diffusion process, measurement error, and IIV are assumed mutually independent for all i, t, and j.

Algorithm

The Extended Kalman Filter (EKF) was implemented to estimate state variables on individual level. The first step of this efficient recursive algorithm is called one-step prediction equations, which are the optimal (minimum variance) prediction of the state mean and covariance at jth measurement, can be calculated by solving the state and state covariance prediction equations from measurement time t_{j-1} until t_j, that is,

$$\frac{d\hat{X}_{i(t|j-1)}}{dt} = g(\hat{X}_{i(t|j-1)}, d_i, \phi_i)$$

$$\frac{d\hat{P}_{i(t|j-1)}}{dt} = A_{it}P_{i(t|j-1)} + P_{i(t|j-1)}A_{it}^T + \sigma_w\sigma_w^T$$

$$A_{it} = \frac{\partial g}{\partial x}\Big|_{x=\hat{x}_{i(t|j-1)}}, t \in [t_{j-1}, t_j]$$

Next, the EFK one-step output prediction equations are calculated by

$$\hat{Y}_{i(t|j-1)} = E[Y_{i(j-1)}, \dots, Y_{i1}] = f(\phi_i, \hat{X}_{i(j|j-1)})$$

$$R_{i(j|j-1)} = Var[Y_{ij}|Y_{i(j-1)}, \dots, Y_{i1}] = C_{ij}P_{i(j|j-1)}C_{ij}^T + \Sigma$$

$$C_{ij} = \frac{\partial f}{\partial x}\Big|_{x=\hat{x}_{i(j|j-1)}}$$

The one-step output prediction $\hat{Y}_{i(t|j-1)}$ is the optimal prediction of the jth measurement before that measurement is taken while $R_{i(j|j-1)}$ is the expected covariance for that prediction. $R_{i(j|j-1)}$ is thus the sum of the state covariance associated with the observed states and the covariance of the actual measurement.

Finally, the one-step state and state covariance predictions are updated by conditioning on the jth measurement using the EKF state update equations, that is,

$$K_{ij} = P_{i|j-1} C_{ij}^T R_{i(j|j-1)}^{-1}$$

$$\hat{X}_{i(j|j)} = \hat{X}_{i(j|j-1)} + K_{ij}(Y_{ij} - \hat{Y}_{i(j|j-1)})$$

$$P_{i(j|j)} = P_{i(j|j-1)} - K_{ij} R_{i(j|j-1)} K_{ij}^T$$

where $\hat{X}_{i(j|j)}$ is the updated state estimate, $P_{i(j|j)}$ is the updated state covariance, and K_{ij} is called the Kalman gain. The optimal state estimate $\hat{X}_{i(j|j)}$ is equal to the best state prediction $\hat{X}_{i(j|j-1)}$ before the measurement is taken plus a correction term consisting of an optimal weighting value times the difference between the measurement Y_{ij} and the one-step prediction of its value.

Results

Raw data visualization

There exist great inter- and intra- individual variability in the observed data in Appendix Figure 3, indicating it's very necessary to develop a hierarchical and stochastic differential equation-based model, respectively.
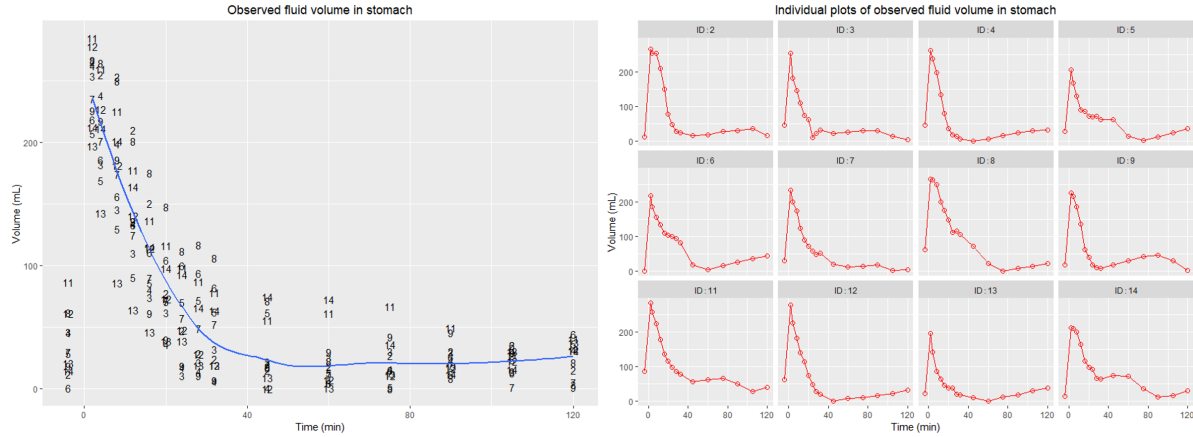
Appendix Figure 3: Great inter- and intra- individual variability were shown in the left and right panel, respectively. Left: Observed fluid volume in the stomach (mL) vs Time (min). The blue curve is the locally weighted smoothing line among different subjects in the time course and the scattered numbers represent the ID number of each subjects. Right: Individual observed fluid volume in the stomach (mL) vs Time (min)

Model Selection

The three SDE models in the Appendix Table 1 showed significant improvement in goodness of fit compared to the traditional ODE model. And the parsimonious model with lowest AIC—SDE with dynamic gastric emptying coefficient model was considered to best characterize the time course of stomach fluid volume and was chosen for further analysis, implying the fluctuation of gastric emptying coefficient might be the driven factor for the system dynamics of fluid volume in the stomach.

Appendix Table 1: A brief summary of the tested models

Models	AIC	Diffusion process
ODE	1416.405	$\alpha = 0$ and $\sigma_w = 0$
SDE with dynamic secretion coefficient	1294.273	$d\log Kse_i(t) = -\alpha(\log Kse_i(t) - \log Kse_i*) + \sigma_{wKse}dW_i(t)$
SDE with dynamic gastric emptying coefficient	1231.567	$d\log Kge_i(t) = -\alpha(\log Kge_i(t) - \log Kge_i*) + \sigma_{wKge}dW_i(t)$
SDE with dynamic secretion and gastric emptying coefficients	1231.567	$d\log Kse_i(t) = -\alpha(\log Kse_i(t) - \log Kse_i*) + \sigma_{wKse}dW_i(t)$ and $d\log Kge_i(t) = -\alpha(\log Kge_i(t) - \log Kge_i*) + \sigma_{wKge}dW_i(t)$

Optimal model results

The estimates of the typical population mean of Kge and Kse are all in line with physiological range in the literature in both SDE and ODE. The proportional and additive measurement error were decreased from 14.4% to 4.10% and 17.6 mL to 4.74 mL for the fluid volume in the stomach compartment by SDE compared with ODE, respectively, indicating a large part of intra-individual variability belong to the system dynamics especially the fluctuation of gastric emptying coefficient Kge. These serial correlated errors would be ignored and empirically assigned to the measurement error if we adopt the traditional ODE method. The relative standard error (RSE) estimates were more or less the same for the SDE and ODE models.

Appendix Table 2: Parameter Estimation of SDE with dynamic gastric emptying rate parameter and ODE

| Model		SDE		ODE	
Parameters		Estimates	RSE(%)	Estimates	RSE(%)
Physiological	Kge (/min)	0.0514	25.1	0.0714	9.01
	Kse (mL/min)	1.52	18.6	1.59	13.9
Intra-individual variability	σ_{wkge}(min$^{-1/2}$)	0.251	13.6	0	na
	α_{kge}(/min)	0.0204	38.9	0	na
	σ_{prop} (%)	3.57	24.6	14.4	20.2
	σ_{add} (mL/min)	4.83	12.4	17.7	8.59
Inter-individual variability	ω_{Kge} (%)	60.6	49.9	27.5	46.2
	ω_{Kse} (%)	18.6	na	21.6	na

RSE = Estimates/SE*100(%), na-not available

The time-varying nature of Kge was shown in the Appendix Table 2 compared to the constant estimation in ODE for each individual in this study. This is the estimation of the hidden state based on observations at each measurement time, which contributes most to the dynamics of the stomach fluid volume system.

The one-step individual SDE predictions, updates, and prediction interval are shown together with the individual ODE predictions in Appendix Figure 4. The observed discrepancy between the SDE and ODE predictions are due to the SDE predictions being conditioned on all previous observations and therefore updated at each sampling time (visualized by the vertical lines in the SDE predictions).

Appendix Figure 4: Estimation of logrithm of the gastric emptying coefficient Kge at each measurement time for each individual based on observed records.
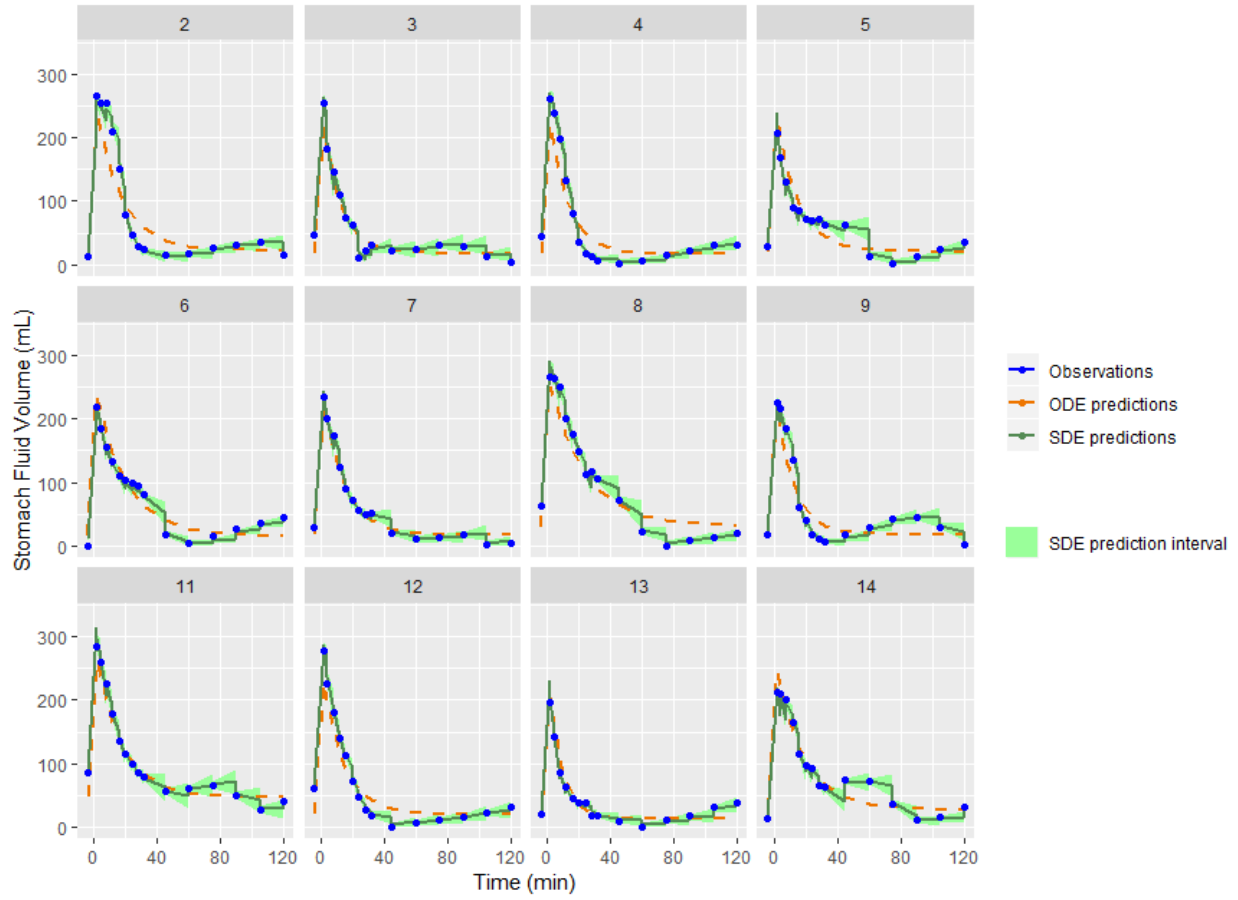
Appendix Figure 5: Individual plots of the one-step SDE predictions, updates, and prediction interval plotted together with the ODE predictions.
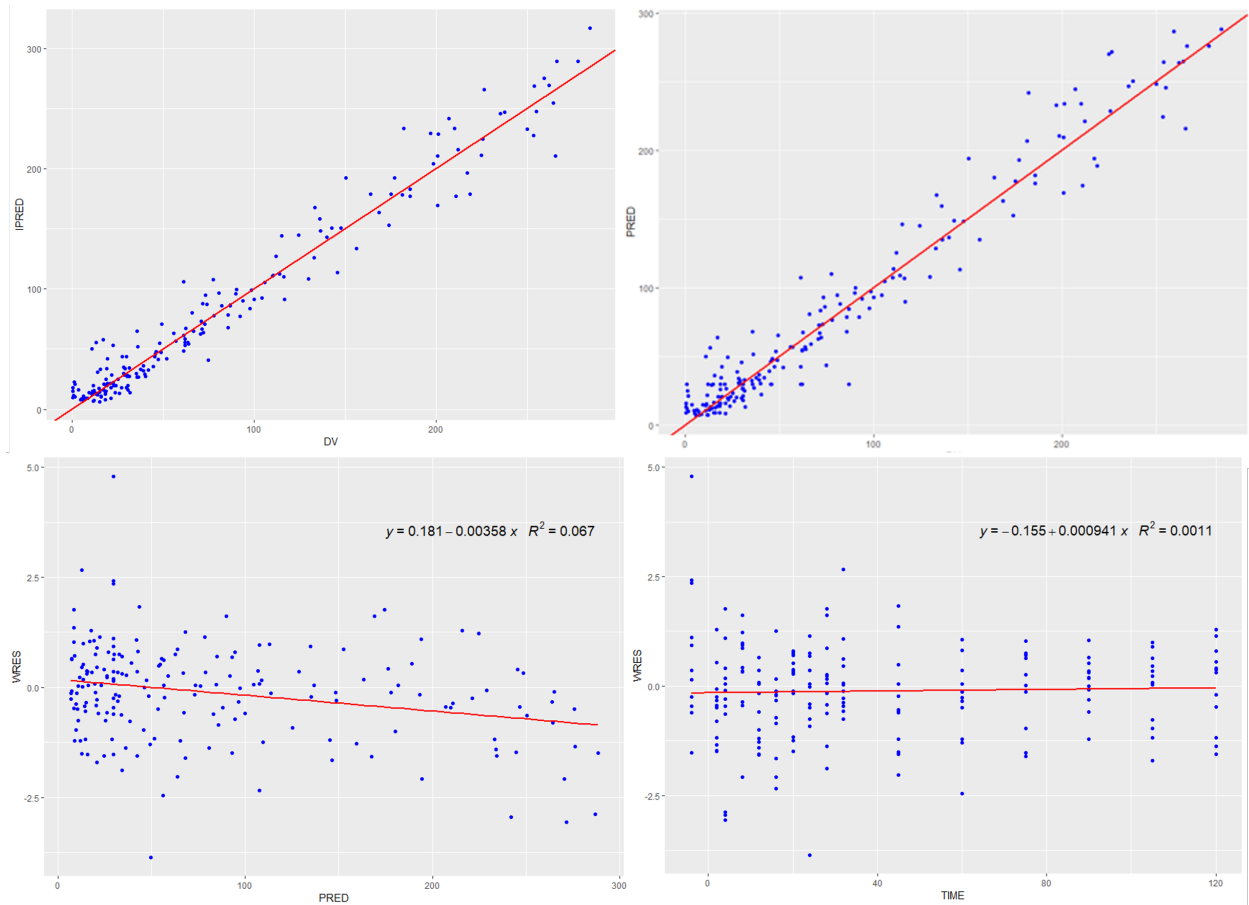
Appendix Figure 6: Goodness-of-fit plots (A-D) and visual predictive check of the SDE with dynamic gastric emptying coefficient based hierarchical model of the fluid volume in the stomach among healthy subjects with EKF estimation algorithm. DV: observed stomach volume (mL). IPRED: individually predicted stomach fluid volume (mL); PRED: population predicted stomach volume (mL); WRES: weighted residual errors.

Optimal model diagnostics

The individual predictions (IPRED) are generated by assuming a new observation based on estimated individual estimated in the population, while the population predictions (PRED) are targeted to the new observations of a new individual based on the estimated parameter distributions. As shown in the A and B plots of Appendix Figure 6, the dots scatter evenly along the two sides of the red line y=x when comparing the prediction and observed fluid volume in the stomach, indicating a good prediction quality. We also investigated on pattern of the weighted residual errors (WRES: $(Y_{ij}-Y_{i(j|j-1)})/R_{j|j-1}$) and found it didn't change as the population prediction

value and time. This means that the WRES is a white noise. Therefore, the selected optimal model was demonstrated to be well developed.

Conclusion and Discussion

The SDE-based mixed effects model developed from our study successfully characterized the volume profiles in the stomach among the healthy subjects and provided a robust approximation of the physiological parameters in the very dynamic system. The remarkable intra-individual variability could be further decomposed into system dynamic terms and measurement error terms in the SDE model instead of empirically assigning all the intra-individual variability to measurement error in the traditional ODE method. The system dynamics were best captured by the random fluctuation of the gastric emptying coefficient.

Project 2: Human Duodenal MMC Phase 3 Motility Model Based on Manometry Readings

Objectives

An insight into phase 3, the decisive factor of gastrointestinal motility and secretion, helps us better predict drug transit and dissolution in gastrointestinal tract and subsequent drug absorption profile. This study is primarily aimed at building up a model of duodenal MMC phase 3 motility behavior in the fasted state based on manometry data via time series analysis.

Data Exploration

The original manometry data was sampled every 0.1s for 7hrs. For the purpose of course project, the data in this study was subsetted by extracting one value out of every 1s during the phase 3 activity of first MMC cycle after determination at mid-duodenum of one subject. The summary info and plot along the time course are displayed below.
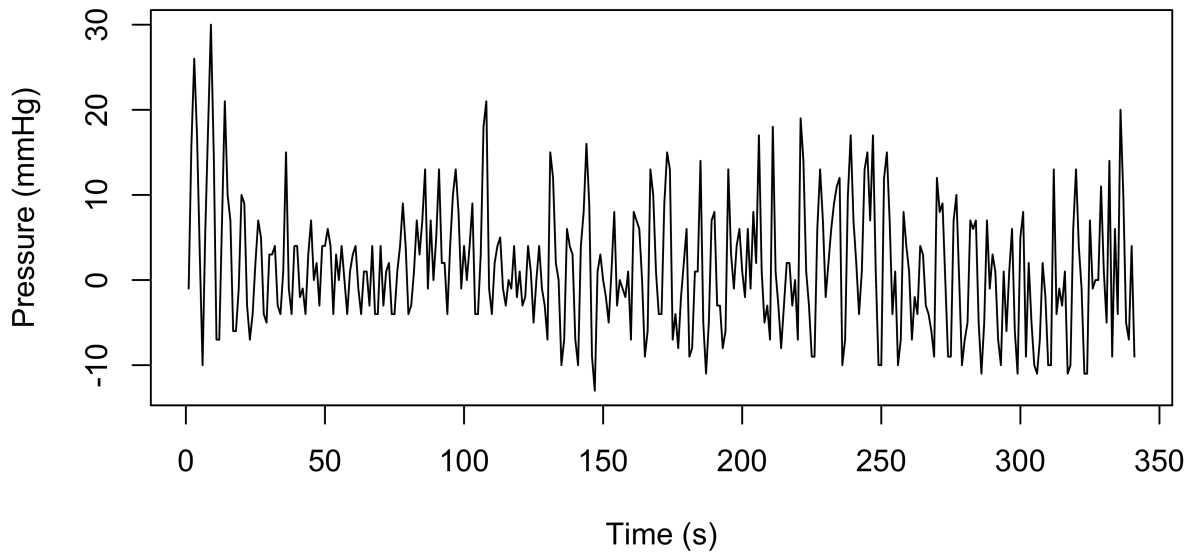

```
##     Min. 1st Qu.  Median    Mean 3rd Qu.     Max.
## -13.000  -4.000   1.000   1.413   6.000   30.000
```

```
## [1] "SD = 7.48"
```

One may first want to check the potential seasonality since phase 3 is known for regular contractions relative to the other two phases. There is an obvious pattern occurring between every 6 lags and the period of the most prominant frequency in the periodogram is also 6s, indicating a mainstream period of 6s per peak in phase 3, which is in concert with the literature claim "10-12 contractions per minute".

Series select2

Smoothed Periodogram

frequency
bandwidth = 0.00704
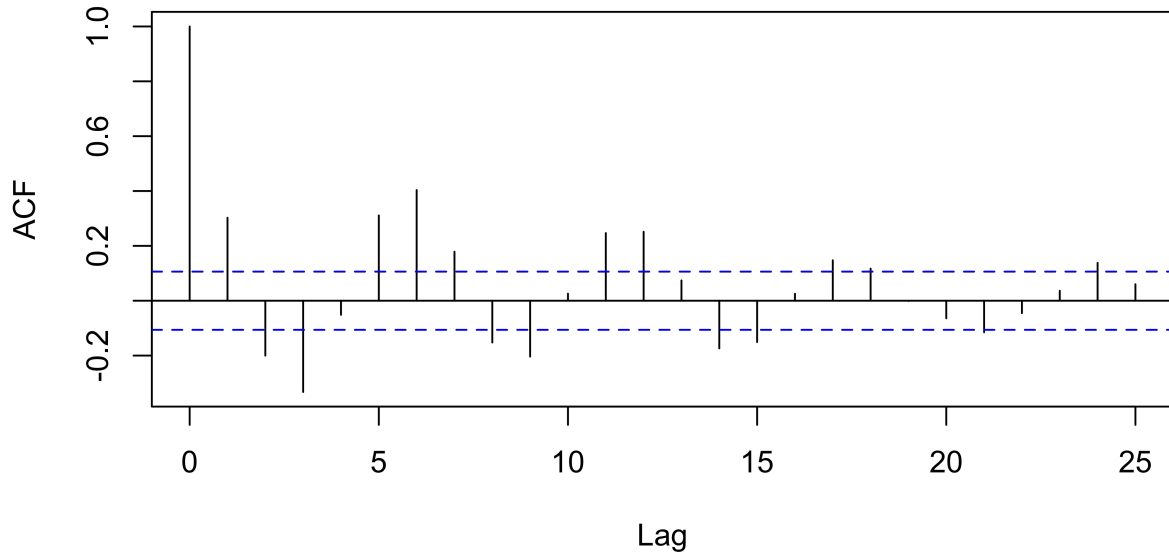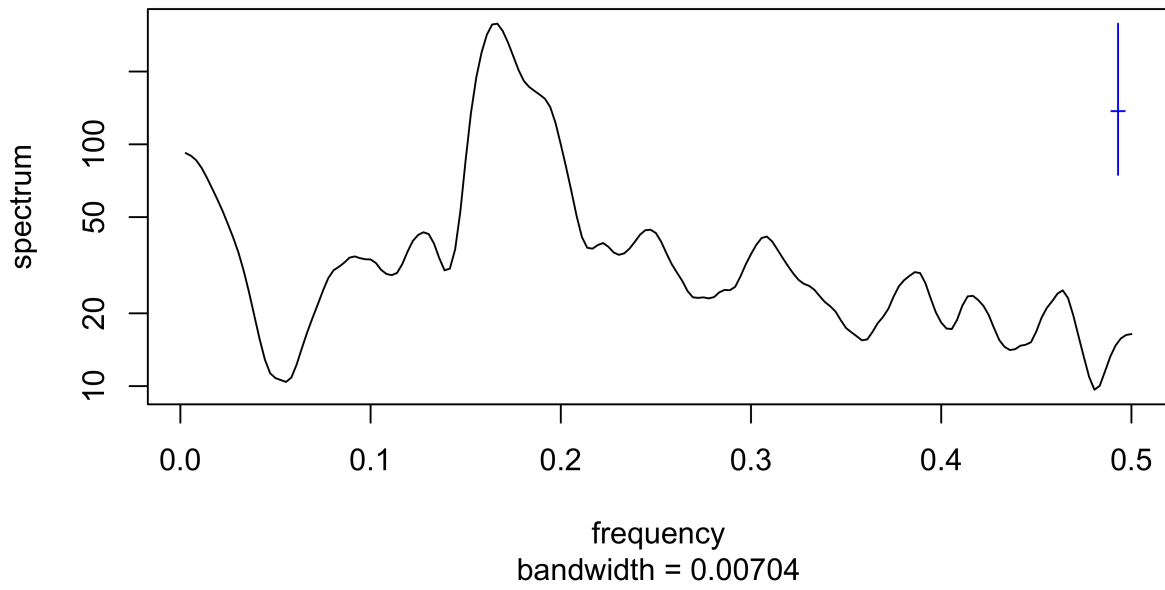
```
## [1] "period = 6"
```

After differencing by 6 lags, the seasonal pattern disappear in the ACF plot and time plot looks more regularized than before, suggesting adding a seasonal component to the model.

Series diff_6
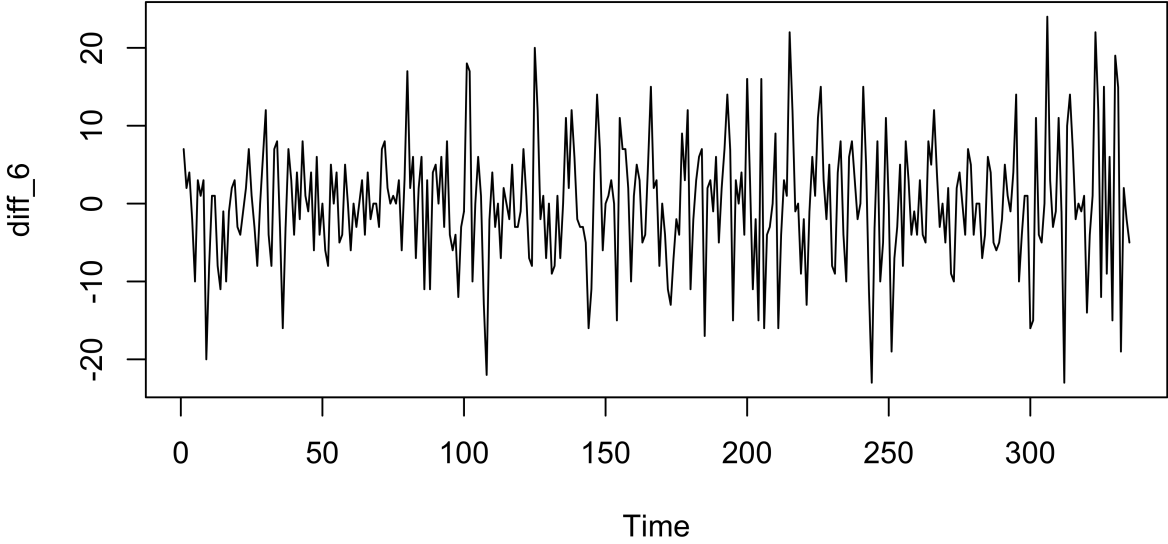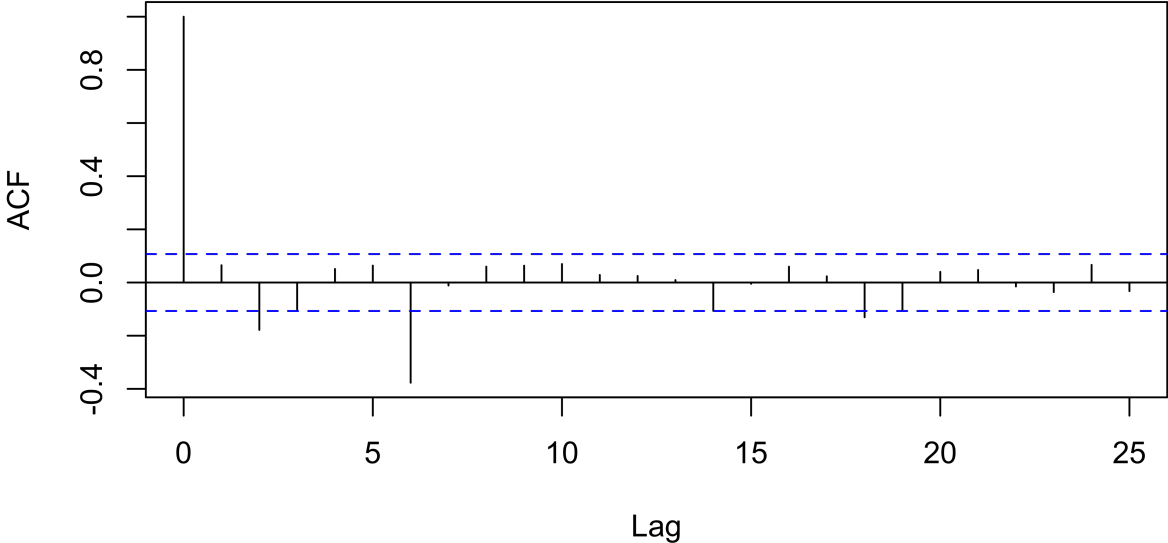
ARMA Model

A time series data can be decomposed into three components: trend, seasonal and random. It's easier to set out to find a suitable autoregressive and moving average (ARMA) model for the random part.

```
ts_1 <- ts(select2,frequency = 6)
de_1 <- decompose(ts_1)
plot(de_1)
```


Decomposition of additive time series

- We seek to fit a stationary Gaussian ARMA(p,q) model with parameter vector $\theta = (\phi_{1:p}, \psi_{1:q}, \mu, \sigma^2)$ given by

$$\phi(B)(Y_n - \mu) = \psi(B)\epsilon_n,$$

  where

$$
\begin{aligned}
\mu &= \mathbb{E}[Y_n] \\
\phi(x) &= 1 - \phi_1 x - \cdots - \phi_p x^p, \\
\psi(x) &= 1 + \psi_1 x + \cdots + \psi_q x^q, \\
\epsilon_n &\sim \text{iid } N[0, \sigma^2].
\end{aligned}
$$

- Let's tabulate some AIC values for a range of different choices of p and q. The ARMA(2,2) model provides a good preditve power. We will choose it for the subsequent complicated model.
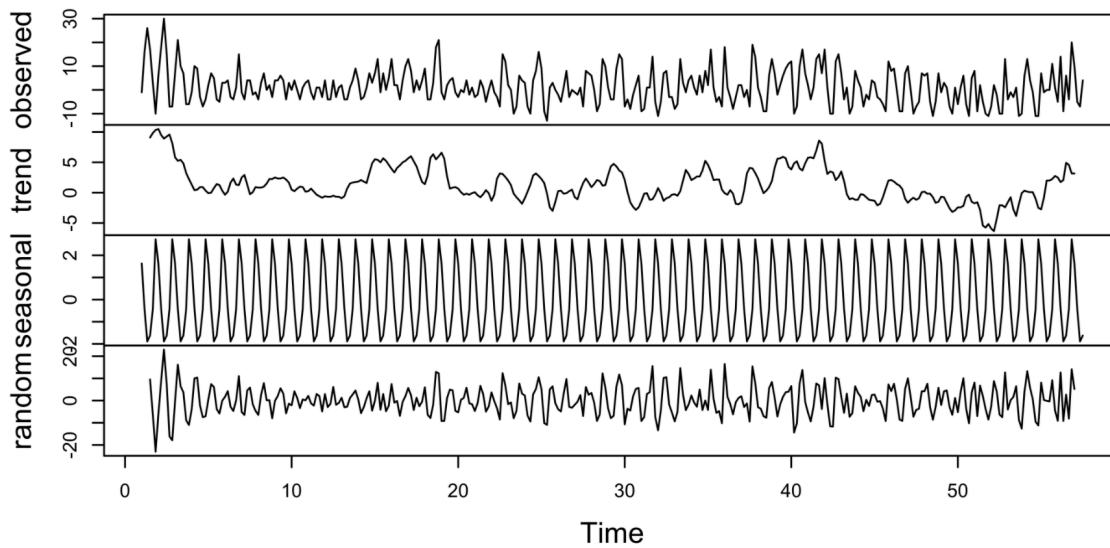
	MA0	MA1	MA2	MA3	MA4	MA5
AR0	2226.57	2209.52	2034.27	2032.62	1915.72	1905.42
AR1	2222.87	2205.10	2035.98	2028.27	1913.81	1902.18
AR2	2124.90	1899.59	1836.89	1838.53	1839.91	1841.32
AR3	2039.88	1858.61	1838.44	1840.57	1842.43	1843.29
AR4	1987.00	1848.36	1839.58	1842.37	1837.83	1844.04

- Let's fit the ARMA(2,2) model recommended by consideration of AIC.

```
arima_1<-arima(rand_1, order = c(2,0,2))
arima_1
```

```
## 
## Call:
## arima(x = rand_1, order = c(2, 0, 2))
## 
## Coefficients:
##          ar1      ar2      ma1     ma2  intercept
##       0.8745  -0.7877  -1.6479  0.6479     -9e-04
## s.e.  0.0396   0.0346   0.0576  0.0572      8e-04
## 
## sigma^2 estimated as 13.2:  log likelihood = -912.45,  aic = 1836.89
```

- In the first plot, residuals generally distributed aroud the line. Thus the residuals follow normal distribution. The Shapiro-Wilk normality test also shows we can not reject the null hypothesis. In the third plot, the residuals generally distributed aroud the line=0. In the fourth plot, residuals are generally inside the dashed line showing pointwise acceptance regions at the 95% confidence level under a null hypothesis of no correlation between noise. Therefore, we can not reject the null hypothesis of Gaussian noise.

Normal Q-Q Plot
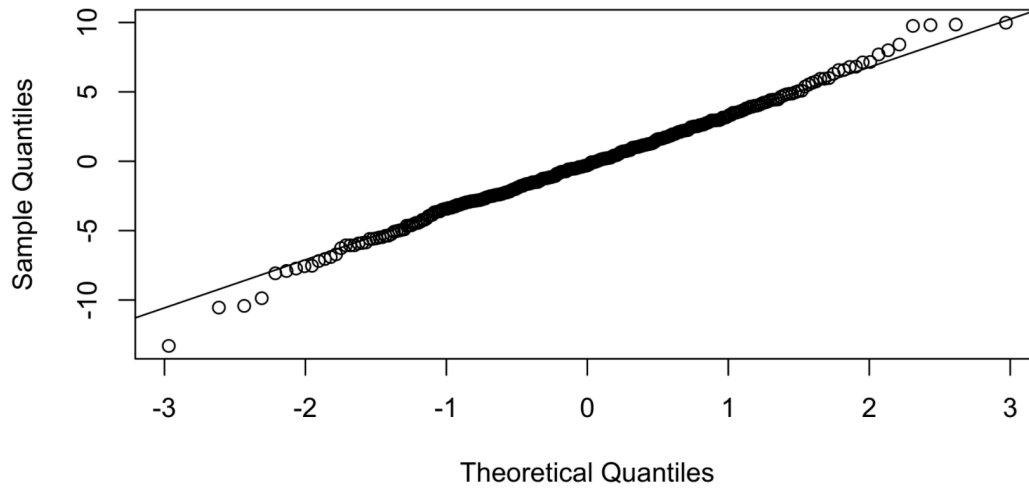

```
##
##  Shapiro-Wilk normality test
##
## data:  arima_1$residuals
## W = 0.99509, p-value = 0.3684
```
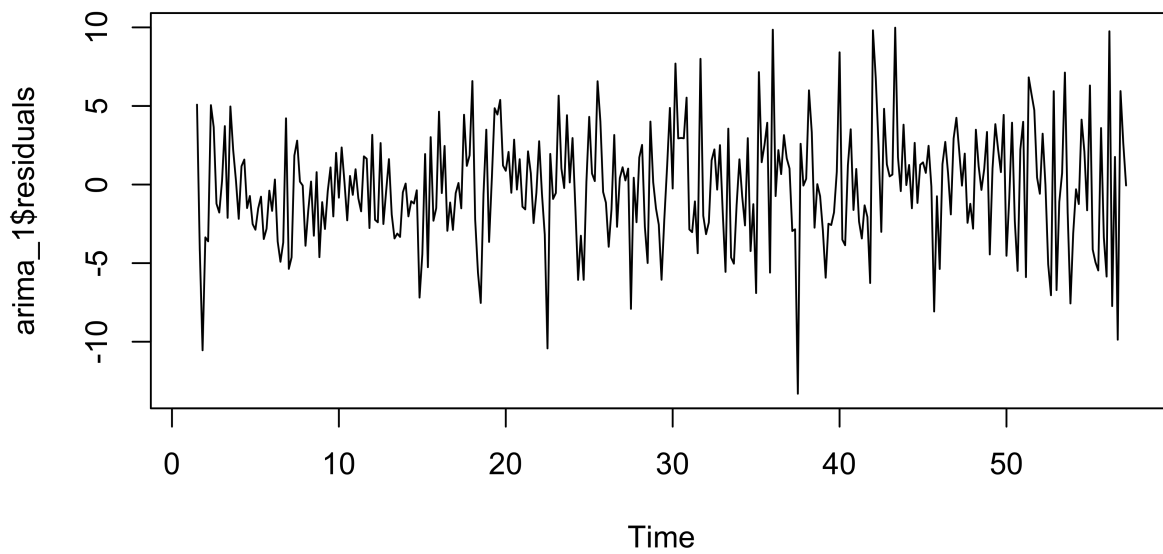
Series arima_1$residuals
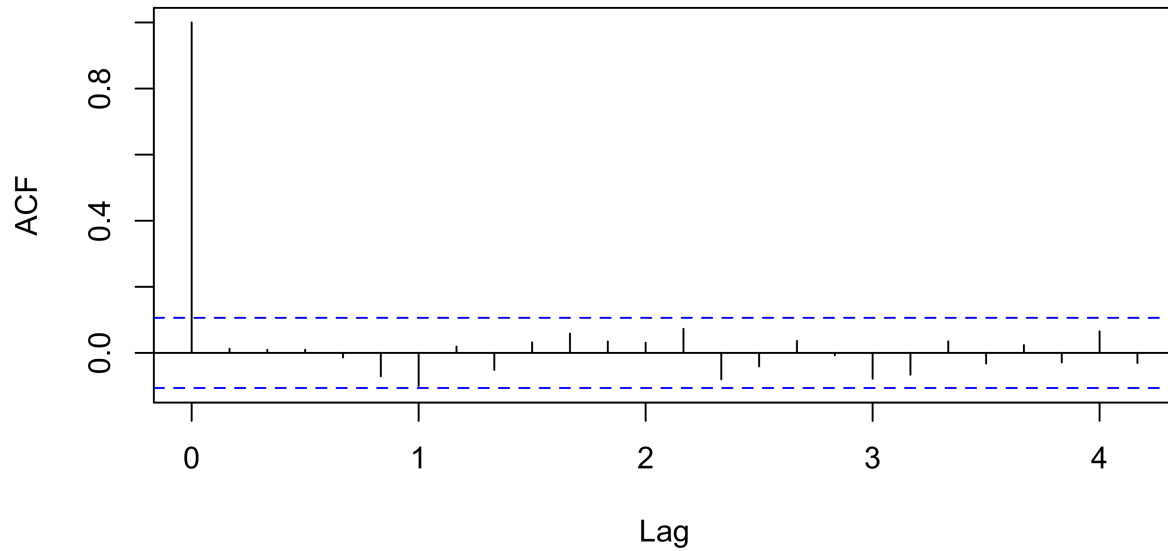
SARMA Model

- Let's continue by adding the seasonable component. The SARMA$(2, 2) \times (P, Q)_6$ model is
$$\phi(B)\Phi(B^6)(Y_n - \mu) = \psi(B)\Psi(B^6)\epsilon_n, \text{ where } \{\epsilon_n\} \text{ is a white noise process and}$$

$$\mu = \mathbb{E}[Y_n]$$
$$\phi(x) = 1 - \phi_1 x - \phi_2 x^2,$$
$$\psi(x) = 1 + \psi_1 x + \psi_2 x^2,$$
$$\Phi(x) = 1 - \Phi_1 x - \cdots - \Phi_p x^P,$$
$$\Psi(x) = 1 + \Psi_1 x + \cdots + \Psi_q x^Q.$$

- We need to tabulate some AIC values for a range of different choices of P and Q. The $SARMA(2, 2) \times (2, 1)_6$ model is chosen based on the AIC table.

	SMA0	SMA1	SMA2	SMA3
SAR0	2231.19	2229.76	2227.34	2229.33
SAR1	2228.80	2227.12	2228.86	2230.44
SAR2	2226.62	2225.54	2227.07	2228.15
SAR3	2228.53	2227.47	2228.48	2227.65

```
sarima_1<-arima(select2,order = c(2,0,2),seasonal=list(order=c(2,0,0),period=6))
sarima_1
```

```
##
## Call:
## arima(x = select2, order = c(2, 0, 2), seasonal = list(order = c(2, 0, 0), period = 6))
##
## Coefficients:
##          ar1      ar2      ma1     ma2    sar1    sar2  intercept
##       0.7825  -0.8281  -0.5906  0.5632  0.1422  0.1296     1.5357
## s.e.  0.0796   0.0759   0.1103  0.1015  0.0662  0.0628     0.4305
##
## sigma^2 estimated as 38.87:  log likelihood = -1105.31,  aic = 2226.62
```

Series sarima_1$residuals
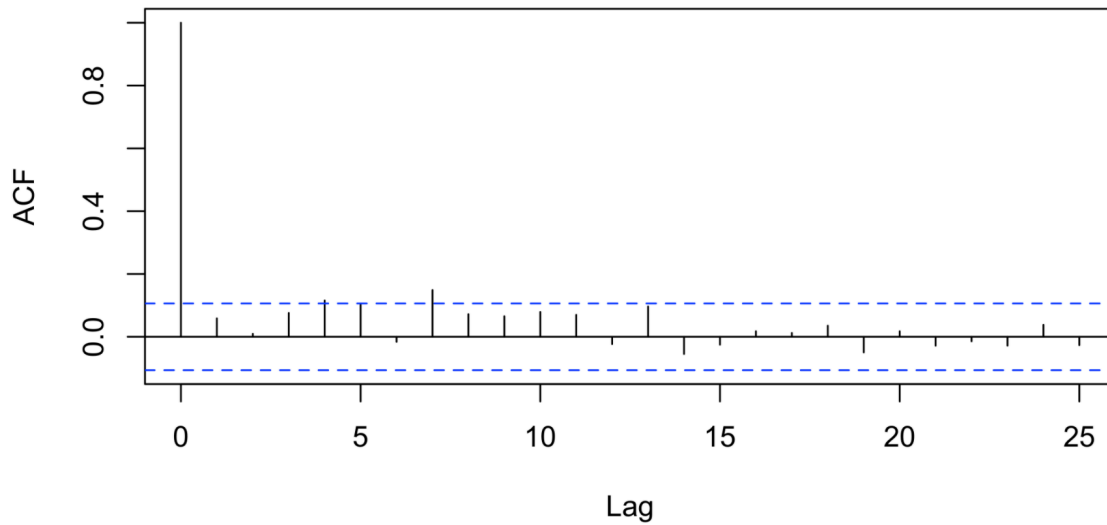
Therefore, the SARIMA model is acceptable.

6. Forecast

- To better examine the model constructed, 90% of the data is taken as train data and 10% as test data to do the forecasting. Most of the points were captured in the 95% confidence interval except some extreme peaks.

```
train <- select2[1:((0.9)*length(select2))]
test <- select2[(0.9*length(select2)+1):length(select2)]
train22 <- arima(train,order = c(2,0,2),seasonal=list(order=c(2,0,0),period=6))
pred_ <- predict(train22, n.ahead = (length(select2)-(0.9*length(select2))))
ts.plot(test,pred_$pred,col=1:2)
u=pred_$pred+1.96*pred_$se
l=pred_$pred-1.96*pred_$se
xx = c(time(u), rev(time(u)))
yy = c(l, rev(u))
polygon(xx, yy, border = 8, col = gray(.6, alpha = .2))
lines(pred_$pred, type="p", col=2)
```
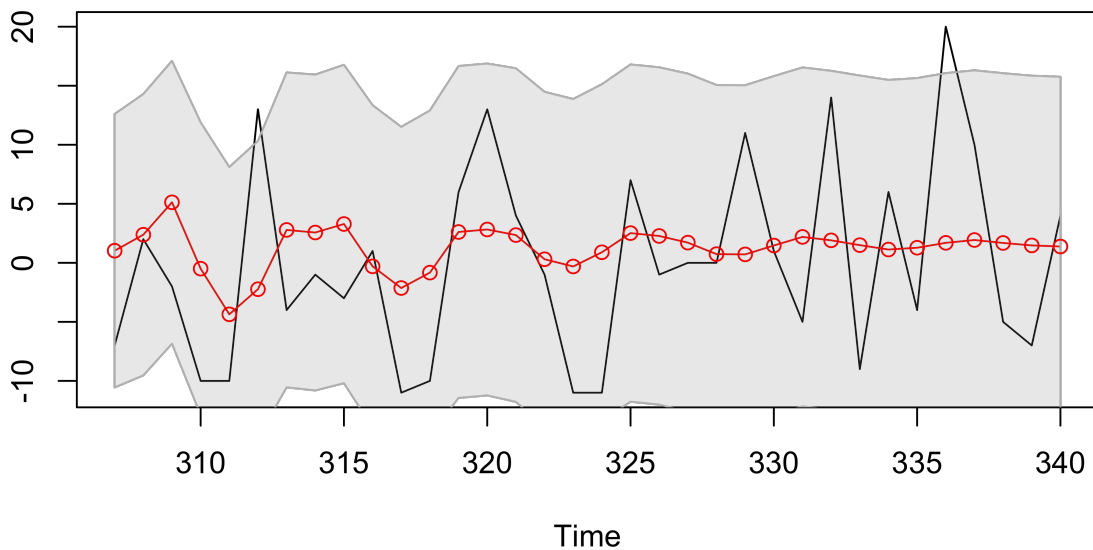

Conclusion and Discussion

1) The SARMA$(2,2)\times(2,0)_6$ model has been successfully constructed to simulate the MMC phase 3 motility pattern, though some further optimization might be needed to account for the extreme peaks;

2) The unsatisfactory part can be explained as the data has an unstable standard deviation and is right skewed. However, many time series analysis methods are built on the assumption

79

of weak stationary: i. constant mean value function; and ii. the autocovariance function only depends on the length of time interval function;

3) In the next step, we will aggregate all the phase 3 manometry data of the subject in mid-duodenum throughout the 7hrs to get a better idea of the motility pattern along the time course accumulate the phase 3 manometry data of all the subjects at different sites along the gastrointestinal tract, build up a hierarchical model to investigate the intra- and inter-subject variance and find out some significant covariable.

Project 3: Human Proximal Small Intestinal pH Model

Introduction

A better characterization of the pH dynamics may help better explain and/or predict the great intra- and inter- individual variabilities in the drug dissolution, solubility and absorption profiles. The SmartPill motility testing system features an ingestible capsule that measures pressure, pH, transit time and temperature as it passes through the entire gastrointestinal tract, making it possible for recording the pH fluctuation in the proximal small intestine. This study is primarily aimed at building up a model reflecting the pH dynamics in the proximal small intestine based on SmartPill data via hidden Markov method.

Data Exploration

The subset individual pH data used in this study starts right after gastric emptying, and ends when the pH reaches a relative stable level. The duration is 30 mins with a sampling frequency of 5s. The summary info and plot along the time course are displayed below. The pH fluctuations are

great in the beginning and dies down gradually as time goes by. The pH statistics of this person all fall into the normal scope[14].

Time after gastric emptying (s)
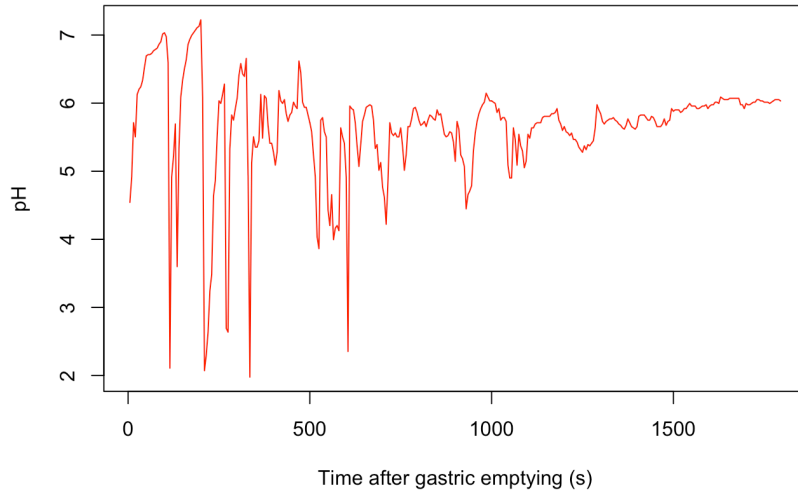
```
##     Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.977   5.505   5.769   5.652   5.995   7.222
```

Model Building

The HCl (pH=1-2) is emptied from the stomach into the duodenum in a pulse packet fluid fashion and is neutralized by HCO3- (pH=7-7.5) secreted from the duodenal cluster unit, especially the pancreas. The intestinal pH is sustained at a reasonable range by the feedback mechanism of human body via adjusting the entry of HCl and HCO3-.[10]
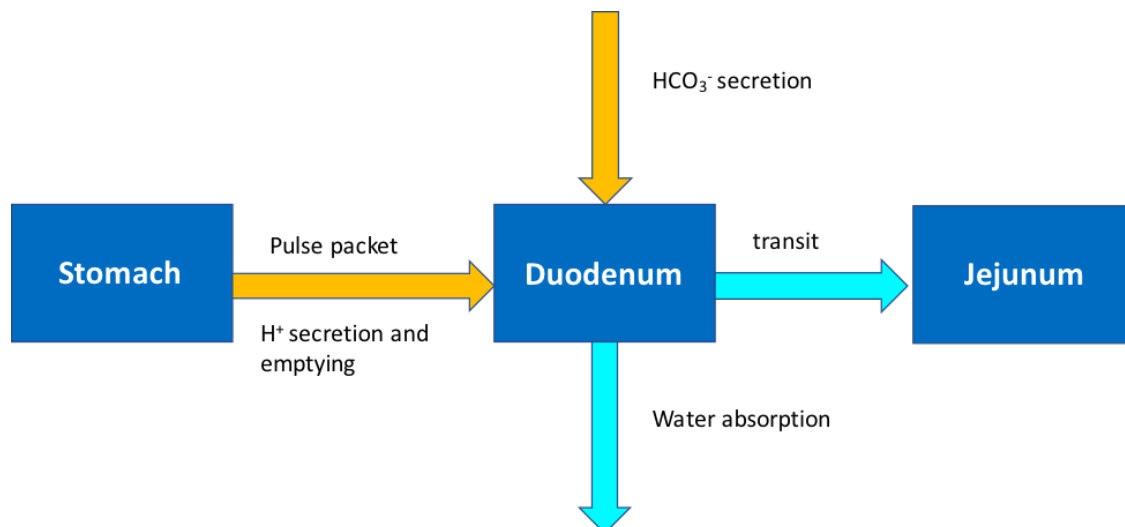
The model proposed in this study is:

$$S_n = S_{n-1} + A * N * B * C,$$
$$if \ S > 0, \ D = -log10(S),$$
$$if \ S < 0, \ D = 7.5 + (log10(-S) * 0.07143)$$

where S denotes the concentration of HCl when it's >0 and the concentration of OH- when it's <0. A can be -1 or 1, and follows a Bernoulli distribution with different probabilities based on the level of S. We can indicate the input from either Stomach or Pancreas by using A. N represent the number of pulses coming from Stomach or Pancreas in a given time period and follows a Poisson process. B refers to the amount of H+ or OH-, and follows a normal distribution. C component means the amount of H+ or OH- every pulse brings is decreasing at an exponential way, as the capsule travels away from the stomach and entry hole of pancreatic secreation. D is state varible of pH. Since the pH of HCO3- falls between 7 to 7.5, I make some restriction on D to make sure it always stay in a reseanable range.

```
sp_statenames <- c("S","D")
sp_paramnames <- c("lambda1","lambda2","mu1","sigma1","sigma2","ub","lb","p1","p2")
sp_obsnames <- colnames(X202108)[2]

sp_test <- c(
  lambda1 = 0.05,
  lambda2 = 0.009,
  mu1 = 0.055,
  sigma1 = 0.01,
  sigma2 = 0.01,
  ub = 0.00000316,
  lb = 0.000000316,
  p1 = 0.001,
  p2 = 0.999
)

sp_rprocess <- "
    double dN = rpois(lambda1*dt);
    if (S>ub){
    S += (2*rbinom(1,p1)-1)*(dN*rnorm(mu1,sigma1)*exp(-lambda2*t));
    }
    else if (lb<S && S<ub){
    S += (2*rbinom(1,0.5)-1)*(dN*rnorm(mu1,sigma1)*exp(-lambda2*t));
    }
    else {
    S += (2*rbinom(1,p2)-1)*(dN*rnorm(mu1,sigma1)*exp(-lambda2*t));
    }
    if (S>0)
    D=-log10(S);
    else
    D=7.5+(log10(-S)*0.07143);
"
```

```
sp_dmeasure <- "
    lik = dnorm(pH,D,sigma2,give_log);
"

sp_rmeasure <- "
    pH = rnorm(D,sigma2);
"

sp_fromEstimationScale <- "
    Tlambda1 = exp(lambda1);
    Tlambda2 = exp(lambda2);
    Tmu1 = exp(mu1);
    Tsigma1 = exp(sigma1);
    Tsigma2 = exp(sigma2);
    Tub = exp(ub);
    Tlb = exp(lb);
    Tp1 = expit(p1);
    Tp2 = expit(p2);
"

sp_toEstimationScale <- "
    Tlambda1 = log(lambda1);
    Tlambda2 = log(lambda2);
    Tmu1 = log(mu1);
    Tsigma1 = log(sigma1);
    Tsigma2 = log(sigma2);
    Tub = log(ub);
    Tlb = log(lb);
    Tp1 = logit(p1);
    Tp2 = logit(p2);
"
```

```
sp_initializer <- "
    S = 0.000006;
    D = 5.22;
"

stopifnot(packageVersion("pomp")>="0.75-1")
sp_pomp <- pomp(
    data=X202108,
    times="Time",
    t0=0,
    rprocess=euler.sim(
      step.fun=Csnippet(sp_rprocess),
      delta.t=5
    ),
    rmeasure=Csnippet(sp_rmeasure),
    dmeasure=Csnippet(sp_dmeasure),
    fromEstimationScale=Csnippet(sp_fromEstimationScale),
    toEstimationScale=Csnippet(sp_toEstimationScale),
    obsnames = sp_obsnames,
    statenames = sp_statenames,
    paramnames = sp_paramnames,
    initializer=Csnippet(sp_initializer)
)

run_level <- 3
switch(run_level,
       {sp_Np=100; sp_Nmif=10; sp_Neval=10; sp_Nglobal=10; sp_Nlocal=10},
       {sp_Np=20000; sp_Nmif=100; sp_Neval=10; sp_Nglobal=10; sp_Nlocal=10},
       {sp_Np=60000; sp_Nmif=300; sp_Neval=10; sp_Nglobal=100; sp_Nlocal=20}
)
```

We would also like to take a look at the simulation results of our pomp model under some tentative parameters.

```
simulate(sp_pomp, params=sp_test, nsim = 10000, states=TRUE) -> x
matplot(time(sp_pomp),t(x["D",1:10,]),type='l',lty=1, xlab="Time",ylab="pH",bty='l',col='blue')
lines(time(sp_pomp),obs(sp_pomp,"pH"),lwd=2,col='black')
```
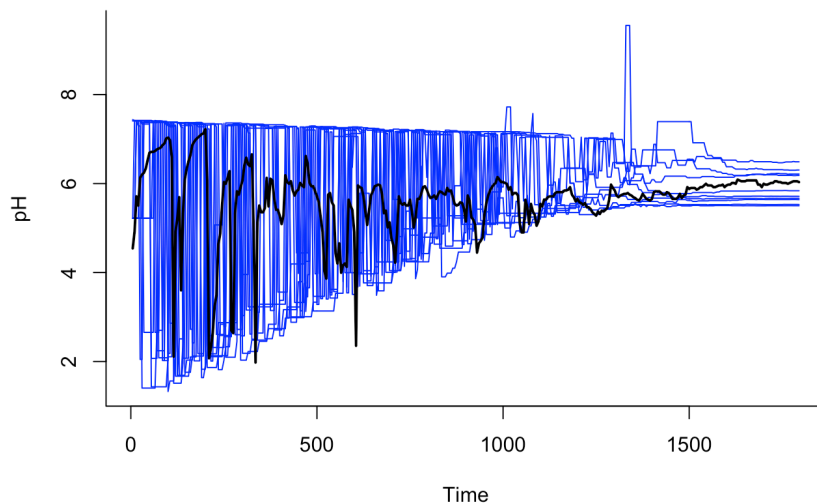
```
pf2 <- replicate(10,pfilter(sp_pomp,Np=5000,params=sp_test))
ll <- sapply(pf2,logLik)
ll
```

```
##  [1] -5382.764 -5818.088 -5503.205 -5261.854 -5553.270 -5338.558 -5580.039
##  [8] -5353.948 -5647.407 -5466.518
```

```
logmeanexp(ll,se=TRUE)
```

```
##                   se
## -5264.15667    69.03386
```

Bibliography

1.   Lennernas H. Human intestinal permeability. *J Pharm Sci* 1998; **87**(4): 403-10.
2.   Yu LX, Lipka E, Crison JR, Amidon GL. Transport approaches to the biopharmaceutical design of oral drug delivery systems: prediction of intestinal absorption. *Adv Drug Deliv Rev* 1996; **19**(3): 359-76.
3.   Waiver of In Vivo Bioavailability and Bioequivalence Studies for Immediate-Release Solid Oral Dosage Forms Based on a Biopharmaceutics Classification System. U.S. Food and Drug Administration; 2017.
4.   M9 Biopharmaceutics Classfication System-based Biowaivers. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; 2019.
5.   Amidon GL, Lennernas H, Shah VP, Crison JR. A Theoretical Basis for a Biopharmaceutic Drug Classification - the Correlation of in-Vitro Drug Product Dissolution and in-Vivo Bioavailability. *Pharmaceut Res* 1995; **12**(3): 413-20.
6.   Amidon GL, Lennernas H, Shah VP, Crison JR. A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. *Pharm Res* 1995; **12**(3): 413-20.
7.   Cannon WB. The movements of the stomach studied by means of the rontgen rays. *Am J Physiol* 1898; **1**: 359-82.
8.   Johnson LR. Physiology of the gastrointestinal tract. 5th ed. London: Academic Press,; 2012. p. 1 online resource (2 volumes).
9.   Ehrlein HJ, Schemann M. The Moving Gut by Hans Jörg Ehrlein and Michael Schemann. http://humanbiology.wzw.tum.de/index.php?id=22&L=1.
10.   Barrett KE. Gastrointestinal physiology. McGraw-Hill's AccessMedicine. 2nd ed. New York, N.Y.: McGraw-Hill Education LLC,; 2014. p. xii, 321 p.
11.   Schubert ML, Peura DA. Control of gastric acid secretion in health and disease. *Gastroenterology* 2008; **134**(7): 1842-60.
12.   Soybel DI. Anatomy and physiology of the stomach. *Surg Clin North Am* 2005; **85**(5): 875-94, v.
13.   Richardson CT, Feldman M. Effects of transdermal scopolamine, alone or in combination with cimetidine, on total 24 hour gastric acid secretion in patients with duodenal ulcer. *Gut* 1986; **27**(12): 1493-7.
14.   Mudie DM, Amidon GL, Amidon GE. Physiological parameters for oral delivery and in vitro testing. *Mol Pharm* 2010; **7**(5): 1388-405.
15.   Doherty GM. Current Diagnosis & Treatment: Surgery. New York: McGraw-Hill; 2014.
16.   Gallagher   N.   Anatomy   &   Physiology   II.   2013. http://anatomyphysiology2.weebly.com/digestive-system.html.
17.   . https://www.britannica.com/science/small-intestine.
18.   Schuster MM, Crowell MD, Koch KL. Schuster atlas of gastrointestinal motility in health and disease. 2nd ed. Hamilton: B.C. Decker; 2002.

19.   Schuster MM, Crowell MD, Koch KL. Schuster Atlas of Gastrointestinal Motility in Health and Disease. 2 ed. London: BC Decker Inc; 2002.

20.   Layer P, Goebell H. [Interdigestive motility and secretion of the gastrointestinal tract]. *Z Gastroenterol* 1987; **25**(12): 769-76.

21.   Jodal M, Eklund S, Sjovall H. Enteric nerves and function of intestinal mucosa. Physiological and pathophysiological aspects. *Dig Dis* 1988; **6**(4): 203-15.

22.   Sjovall H. Meaningful or redundant complexity - mechanisms behind cyclic changes in gastroduodenal pH in the fasting state. *Acta Physiol (Oxf)* 2011; **201**(1): 127-31.

23.   Layer P, Chan AT, Go VL, DiMagno EP. Human pancreatic secretion during phase II antral motility of the interdigestive cycle. *Am J Physiol* 1988; **254**(2 Pt 1): G249-53.

24.   Traynor OJ, Dozois RR, DiMagno EP. Canine interdigestive and postprandial gallbladder motility and emptying. *Am J Physiol* 1984; **246**(4 Pt 1): G426-32.

25.   DiMagno EP. Regulation of interdigestive gastrointestinal motility and secretion. *Digestion* 1997; **58 Suppl 1**: 53-5.

26.   Steingoetter A, Fox M, Treier R, et al. Effects of posture on the physiology of gastric emptying: a magnetic resonance imaging study. *Scand J Gastroenterol* 2006; **41**(10): 1155-64.

27.   Ramsbottom N, Knox MT, Hunt JN. Gastric-Emptying of Barium-Sulfate Suspension Compared with That of Water. *Gut* 1977; **18**(7): 541-2.

28.   Umenai T, Arai N, Chihara E. Effect of the preliminary hydration on gastric emptying time for water in healthy volunteers. *Acta Anaesthesiol Scand* 2009; **53**(2): 223-6.

29.   Wright J, Adams V, Hykin J, et al. The measurement of gastric motor function and transit in man by echo planar magnetic resonance imaging. *Magn Reson Mater Phy* 1994; **2**(3): 467-9.

30.   Oberle RL, Chen TS, Lloyd C, et al. The influence of the interdigestive migrating myoelectric complex on the gastric emptying of liquids. *Gastroenterology* 1990; **99**(5): 1275-82.

31.   Mudie DM, Murray K, Hoad CL, et al. Quantification of gastrointestinal liquid volumes and distribution following a 240 mL dose of water in the fasted state. *Mol Pharm* 2014; **11**(9): 3039-47.

32.   Paixao P, Bermejo M, Hens B, et al. Gastric emptying and intestinal appearance of nonabsorbable drugs phenol red and paromomycin in human subjects: A multi-compartment stomach approach. *Eur J Pharm Biopharm* 2018; **129**: 162-74.

33.   Schiller C, Frohlich CP, Giessmann T, et al. Intestinal fluid volumes and transit of dosage forms as assessed by magnetic resonance imaging. *Aliment Pharm Therap* 2005; **22**(10): 971-9.

34.   Koziolek M, Alcaro S, Augustijns P, et al. The mechanisms of pharmacokinetic food-drug interactions - A perspective from the UNGAP group. *European Journal of Pharmaceutical Sciences* 2019; **134**: 31-59.

35.   Ouyang A, Sunshine AG, Reynolds JC. Caloric Content of a Meal Affects Duration but Not Contractile Pattern of Duodenal Motility in Man. *Digest Dis Sci* 1989; **34**(4): 528-36.

36.   Food-Effect Bioavailability and Fed Bioequivalence Studies. U.S. Food and Drug Administration; 2018.

37.   Abuhelwa AY, Foster DJR, Upton RN. A Quantitative Review and Meta-models of the Variability and Factors Affecting Oral Drug Absorption-Part II: Gastrointestinal Transit Time. *Aaps Journal* 2016; **18**(5): 1322-33.

38.   Jiang Y, Milavetz G, James MO, An GH. A Mechanism-Based Pharmacokinetic Enzyme Turnover Model for Dichloroacetic Acid Autoinhibition in Rats. *J Pharm Sci-Us* 2017; **106**(5): 1396-404.

39.   Cosson VF, Fuseau E. Mixed effect modeling of sumatriptan pharmacokinetics during drug development: II. From healthy subjects to phase 2 dose ranging in patients. *J Pharmacokinet Biop* 1999; **27**(2): 149-71.

40.   Rietbrock S, Merz PG, Fuhr U, et al. Absorption Behavior of Sulpiride Described Using Weibull Functions. *Int J Clin Pharm Th* 1995; **33**(5): 299-303.

41.   Locatelli I, Mrhar A, Bogataj M. Gastric Emptying of Pellets under Fasting Conditions: A Mathematical Model. *Pharmaceut Res* 2009; **26**(7): 1607-17.

42.   Siegel JA, Urbain JL, Adler LP, et al. Biphasic Nature of Gastric-Emptying. *Gut* 1988; **29**(1): 85-9.

43.   Wu K, Cohen EE, House LK, et al. Nonlinear population pharmacokinetics of sirolimus in patients with advanced cancer. *CPT Pharmacometrics Syst Pharmacol* 2012; **1**: e17.

44.   Higaki K, Yamashita S, Amidon GL. Time-dependent oral absorption models. *J Pharmacokinet Pharmacodyn* 2001; **28**(2): 109-28.

45.   Csajka C, Drover D, Verotta D. The use of a sum of inverse Gaussian functions to describe the absorption profile of drugs exhibiting complex absorption. *Pharmaceut Res* 2005; **22**(8): 1227-35.

46.   Stubbs DF. Models of gastric emptying. *Gut* 1977; **18**(3): 202-7.

47.   Dressman JB, Fleisher D, Amidon GL. Physicochemical model for dose-dependent drug absorption. *J Pharm Sci* 1984; **73**(9): 1274-9.

48.   Luner PE, Amidon GL. Description and simulation of a multiple mixing tank model to predict the effect of bile sequestrants on bile salt excretion. *J Pharm Sci* 1993; **82**(3): 311-8.

49.   Yu LX, Amidon GL. A compartmental absorption and transit model for estimating oral drug absorption. *Int J Pharmaceut* 1999; **186**(2): 119-25.

50.   Agoram B, Woltosz WS, Bolger MB. Predicting the impact of physiological and biochemical processes on oral drug bioavailability. *Adv Drug Deliv Rev* 2001; **50 Suppl 1**: S41-67.

51.   Langguth P, Lee KM, Spahn-Langguth H, Amidon GL. Variable gastric emptying and discontinuities in drug absorption profiles: dependence of rates and extent of cimetidine absorption on motility phase and pH. *Biopharm Drug Dispos* 1994; **15**(9): 719-46.

52.   Talattof A, Price JC, Amidon GL. Gastrointestinal Motility Variation and Implications for Plasma Level Variation: Oral Drug Products. *Mol Pharm* 2016; **13**(2): 557-67.

53.   Karatza E, Karalis V. Investigating the Impact of Gastric Emptying on Pharmacokinetic Parameters Using Delay Differential Equations and Principal Component Analysis. *Eur J Drug Metab Ph* 2021.

54.   Henin E, Bergstrand M, Weitschies W, Karlsson MO. Meta-analysis of Magnetic Marker Monitoring Data to Characterize the Movement of Single Unit Dosage Forms Though the Gastrointestinal Tract Under Fed and Fasting Conditions. *Pharmaceut Res* 2016; **33**(3): 751-62.

55.   Yokrattanasak J, De Gaetano A, Panunzi S, Satiracoo P, Lawton WM, Lenbury Y. A Simple, Realistic Stochastic Model of Gastric Emptying. *PLoS One* 2016; **11**(4): e0153297.

56.   Talattof A, Amidon GL. Pulse Packet Stochastic Model for Gastric Emptying in the Fasted State: A Physiological Approach. *Mol Pharmaceut* 2018; **15**(6): 2107-15.

57.   Hens B, Tsume Y, Bermejo M, et al. Low Buffer Capacity and Alternating Motility along the Human Gastrointestinal Tract: Implications for in Vivo Dissolution and Absorption of Ionizable Drugs. *Mol Pharm* 2017.

58. Kalantzi L, Goumas K, Kalioras V, Abrahamsson B, Dressman JB, Reppas C. Characterization of the human upper gastrointestinal contents under conditions simulating bioavailability/bioequivalence studies. *Pharmaceut Res* 2006; **23**(1): 165-76.

59. Clarysse S, Tack J, Lammert F, Duchateau G, Reppas C, Augustijns P. Postprandial evolution in composition and characteristics of human duodenal fluids in different nutritional states. *J Pharm Sci* 2009; **98**(3): 1177-92.

60. Kaunitz JD, Akiba Y. Review article: duodenal bicarbonate - mucosal protection, luminal chemosensing and acid-base balance. *Aliment Pharmacol Ther* 2006; **24 Suppl 4**: 169-76.

61. Hogan DL, Ainsworth MA, Isenberg JI. Review article: gastroduodenal bicarbonate secretion. *Aliment Pharmacol Ther* 1994; **8**(5): 475-88.

62. Fuchs A, Dressman JB. Composition and physicochemical properties of fasted-state human duodenal and jejunal fluid: a critical evaluation of the available data. *J Pharm Sci* 2014; **103**(11): 3398-411.

63. McNamara DP, Whitney KM, Goss SL. Use of a physiologic bicarbonate buffer system for dissolution characterization of ionizable drugs. *Pharmaceut Res* 2003; **20**(10): 1641-6.

64. Koenigsknecht MJ, Baker JR, Wen B, et al. In Vivo Dissolution and Systemic Absorption of Immediate Release Ibuprofen in Human Gastrointestinal Tract under Fed and Fasted Conditions. *Mol Pharmaceut* 2017; **14**(12): 4295-304.

65. Cheng LS, Wong H. Food Effects on Oral Drug Absorption: Application of Physiologically-Based Pharmacokinetic Modeling as a Predictive Tool. *Pharmaceutics* 2020; **12**(7).

66. Abuhelwa AY, Foster DJR, Upton RN. A Quantitative Review and Meta-Models of the Variability and Factors Affecting Oral Drug Absorption-Part I: Gastrointestinal pH. *Aaps Journal* 2016; **18**(5): 1309-21.

67. Drugs@FDA. Silver Spring, MD: U.S. Food and Drug Administration.

68. CellCept (mycophenolate mofetil) [package insert]. San Francisco, CA; F. Hoffmann-La Roche AG; 1995.

69. Armstrong VW, Tenderich G, Shipkova M, et al. Pharmacokinetics and bioavailability of mycophenolic acid after intravenous administration and oral administration of mycophenolate mofetil to heart transplant recipients. *Ther Drug Monit* 2005; **27**(3): 315-21.

70. Bullingham R, Monroe S, Nicholls A, Hale M. Pharmacokinetics and bioavailability of mycophenolate mofetil in healthy subjects after single-dose oral and intravenous administration. *Journal of Clinical Pharmacology* 1996; **36**(4): 315-24.

71. Staatz CE, Tett SE. Clinical pharmacokinetics and pharmacodynamics of mycophenolate in solid organ transplant recipients. *Clin Pharmacokinet* 2007; **46**(1): 13-58.

72. Roberts MS, Magnusson BM, Burczynski FJ, Weiss M. Enterohepatic circulation - Physiological, pharmacokinetic and clinical implications. *Clin Pharmacokinet* 2002; **41**(10): 751-90.

73. Picard N, Cresteil T, Premaud A, Marquet P. Characterization of a phase 1 metabolite of mycophenolic acid produced by CYP3A4/5. *Ther Drug Monit* 2004; **26**(6): 600-8.

74. Naesens M, Kuypers DRJ, Verbeke K, Vanrenterghem Y. Multidrug resistance protein 2 genetic polymorphisms influence mycophenolic acid exposure in renal allograft recipients. *Transplantation* 2006; **82**(8): 1074-84.

75. Blume HH, Midha KK. Bio-International-92, Conference on Bioavailability, Bioequivalence, and Pharmacokinetic Studies. *J Pharm Sci-Us* 1993; **82**(11): 1186-9.

76.   Shah VP, Yacobi A, Barr WH, et al. Evaluation of orally administered highly variable drugs and drug formulations. *Pharmaceut Res* 1996; **13**(11): 1590-4.

77.   Westlake WJ. Bioequivalence Testing - a Need to Rethink - Reply. *Biometrics* 1981; **37**(3): 591-3.

78.   Schuirmann DJ. A Comparison of the 2 One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability. *J Pharmacokinet Biop* 1987; **15**(6): 657-80.

79.   Davit BM, Chen ML, Conner DP, et al. Implementation of a reference-scaled average bioequivalence approach for highly variable generic drug products by the US Food and Drug Administration. *AAPS J* 2012; **14**(4): 915-24.

80.   Patel S, Chauhan V, Mandal J, et al. Single-dose, Two-way Crossover, Bioequivalence Study of Mycophenolate Mofetil 500 mg Tablet Under Fasting Conditions in Healthy Male Subjects. *Clinical Therapeutics* 2011; **33**(3): 378-90.

81.   Almeida S, Filipe A, Neves R, et al. Mycophenolate Mofetil 500-mg Tablet Under Fasting Conditions: Single-Dose, Randomized-Sequence, Open-Label, Four-Way Replicate Crossover, Bioequivalence Study in Healthy Subjects. *Clinical Therapeutics* 2010; **32**(3): 556-74.

82.   Funaki T. Enterohepatic circulation model for population pharmacokinetic analysis. *Journal of Pharmacy and Pharmacology* 1999; **51**(10): 1143-8.

83.   Payen S, Zhang D, Maisin A, et al. Population pharmacokinetics of mycophenolic acid in kidney transplant pediatric and adolescent patients. *Ther Drug Monit* 2005; **27**(3): 378-88.

84.   Cremers S, Schoemaker R, Scholten E, et al. Characterizing the role of enterohepatic recycling in the interactions between mycophenolate mofetil and calcineurin inhibitors in renal transplant patients by pharmacokinetic modelling. *Brit J Clin Pharmaco* 2005; **60**(3): 249-56.

85.   Jiao Z, Ding JJ, Shen J, et al. Population pharmacokinetic modelling for enterohepatic circulation of mycophenolic acid in healthy Chinese and the influence of polymorphisms in UGT1A9. *Brit J Clin Pharmaco* 2008; **65**(6): 893-907.

86.   Yau WP, Vathsala A, Lou HX, Zhou SF, Chan E. Mechanism-Based Enterohepatic Circulation Model of Mycophenolic Acid and Its Glucuronide Metabolite: Assessment of Impact of Cyclosporine Dose in Asian Renal Transplant Patients. *Journal of Clinical Pharmacology* 2009; **49**(6): 684-99.

87.   Sam WJ, Akhlaghi F, Rosenbaum SE. Population Pharmacokinetics of Mycophenolic Acid and Its 2 Glucuronidated Metabolites in Kidney Transplant Recipients. *Journal of Clinical Pharmacology* 2009; **49**(2): 185-95.

88.   de Winter BCM, Neumann I, van Hest RM, van Gelder T, Mathot RAA. Limited Sampling Strategies for Therapeutic Drug Monitoring of Mycophenolate Mofetil Therapy in Patients With Autoimmune Disease. *Ther Drug Monit* 2009; **31**(3): 382-90.

89.   Lennernas H. Human intestinal permeability. *J Pharm Sci-Us* 1998; **87**(4): 403-10.

90.   Yu LX, Lipka E, Crison JR, Amidon GL. Transport approaches to the biopharmaceutical design of oral drug delivery systems: Prediction of intestinal absorption. *Adv Drug Deliver Rev* 1996; **19**(3): 359-76.

91.   Scheubel E, Adamy L, Cardot JM. Mycophenolate Mofetil: Use of a Simple Dissolution Technique to Assess Generic Formulation Differences. *Dissolut Technol* 2012; **19**(1): 52-8.

92.   Bermejo M, Paixao P, Hens B, et al. Linking the Gastrointestinal Behavior of Ibuprofen with the Systemic Exposure between and within Humans-Part 1: Fasted State Conditions. *Mol Pharmaceut* 2018; **15**(12): 5454-67.

93.   Tsume Y, Langguth P, Garcia-Arieta A, Amidon GL. In silico prediction of drug dissolution and absorption with variation in intestinal pH for BCS class II weak acid drugs: ibuprofen and ketoprofen. *Biopharmaceutics & Drug Disposition* 2012; **33**(7): 366-77.

94.   Jantratid E, Janssen N, Reppas C, Dressman JB. Dissolution media simulating conditions in the proximal human gastrointestinal tract: An update. *Pharmaceut Res* 2008; **25**(7): 1663-76.

95.   The United States Phamacopedia USP 31. Rockville: The United States Pharmacopedial Convention Inc; 2008.

96.   Sherwin CMT, Fukuda T, Brunner HI, Goebel J, Vinks AA. The Evolution of Population Pharmacokinetic Models to Describe the Enterohepatic Recycling of Mycophenolic Acid in Solid Organ Transplantation and Autoimmune Disease. *Clin Pharmacokinet* 2011; **50**(1): 1-24.

97.   Pustejovsky J, Stubbs A. Natural Language Annotation for Machine Learning. 1 ed. 1005 Gravenstein Highway North, Sebastopol, CA 95472: O'Reilly Media, Inc.; 2012.

98.   Reed J. NATURAL LANGUAGE PROCESSING – How Life Sciences Companies Are Leveraging NLP From Molecule to Market. 2020. https://drug-dev.com/natural-language-processing-how-life-sciences-companies-are-leveraging-nlp-from-molecule-to-market/.

99.   Kristensen NR, Madsen H, Jorgensen SB. Parameter estimation in stochastic grey-box models. *Automatica* 2004; **40**(2): 225-37.