



## ชื่อบทความ/ผลงาน Data analytics and precision medicine: an overview

ชื่อผู้เขียนบทความ ดร.ภก.วริทธิ์ เรืองเลิศบุญ และ ดร.ภญ.สุพัตรา ลิ้มสุวรรณโชติ

---

### 1. บทคัดย่อ

Precision medicine หรือการแพทย์แม่นยำเป็นกรอบแนวคิดเชิงปฏิบัติที่เข้ามามีบทบาทมากขึ้นในปัจจุบัน โดยเฉพาะในส่วนของ การรักษาโรคที่มีความเสี่ยงสูงต่อการเสียชีวิตหรือกลุ่มโรคที่ยากต่อการรักษา โดย precision medicine ผสมรวมปัจจัยของความแตกต่างระหว่างบุคคลในมิติเชิงพันธุกรรม สภาพแวดล้อม และรูปแบบการใช้ชีวิตของแต่ละบุคคล เพื่อนำมาประกอบการตัดสินใจหรือปรับปรุงแนวทางการรักษาทางคลินิก ปัจจุบันความก้าวหน้าทางวิทยาการข้อมูลเข้ามามีบทบาทส่งเสริมงานวิจัยทางด้าน precision medicine มากขึ้นทั้งในส่วนของ การพัฒนาความรู้เชิงชีววิทยาระดับโมเลกุล (molecular biology) จนถึงการประยุกต์ใช้ทางคลินิก เช่น การสร้างตัวแบบจำลองของยาด้วยโปรแกรมคอมพิวเตอร์เพื่อศึกษาพฤติกรรมทางเภสัชจลนศาสตร์ในสภาวะจำเพาะต่างๆ การประยุกต์ศาสตร์ด้านวิทยาการข้อมูลมีส่วนส่งเสริมต่อความก้าวหน้าของ precision medicine อย่างมีนัยสำคัญ การรู้จักใช้วิทยาการข้อมูลผนวกเข้ากับเภสัชกรรมปฏิบัติจะเป็นการช่วยส่งเสริมให้การใช้ยามีประสิทธิภาพมากขึ้นและเป็นการสร้างองค์ความรู้ใหม่จากฐานข้อมูลที่มีอยู่ ถือเป็น การส่งเสริมบทบาทและภาพลักษณ์ของเภสัชกรในด้านผู้เชี่ยวชาญเกี่ยวกับยาในระบบสาธารณสุขมากขึ้น

### 2. คำสำคัญ (Key words) ของบทความ ( 3 - 5 คำ)

Precision medicine, Personalised medicine, Clinical epidemiology, Physiologically-based pharmacokinetics modelling, Data science

### 3. บทความทางวิชาการฉบับเต็ม

#### 3.1 วัตถุประสงค์เชิงพฤติกรรม

1. เข้าใจความหมายและแนวคิดของการแพทย์แม่นยำ

2. เข้าใจภาพรวมของการนำวิทยาการคอมพิวเตอร์มาประยุกต์เข้ากับศาสตร์ด้านเภสัชวิทยาเชิงคลินิก

### 3.2 เนื้อหา

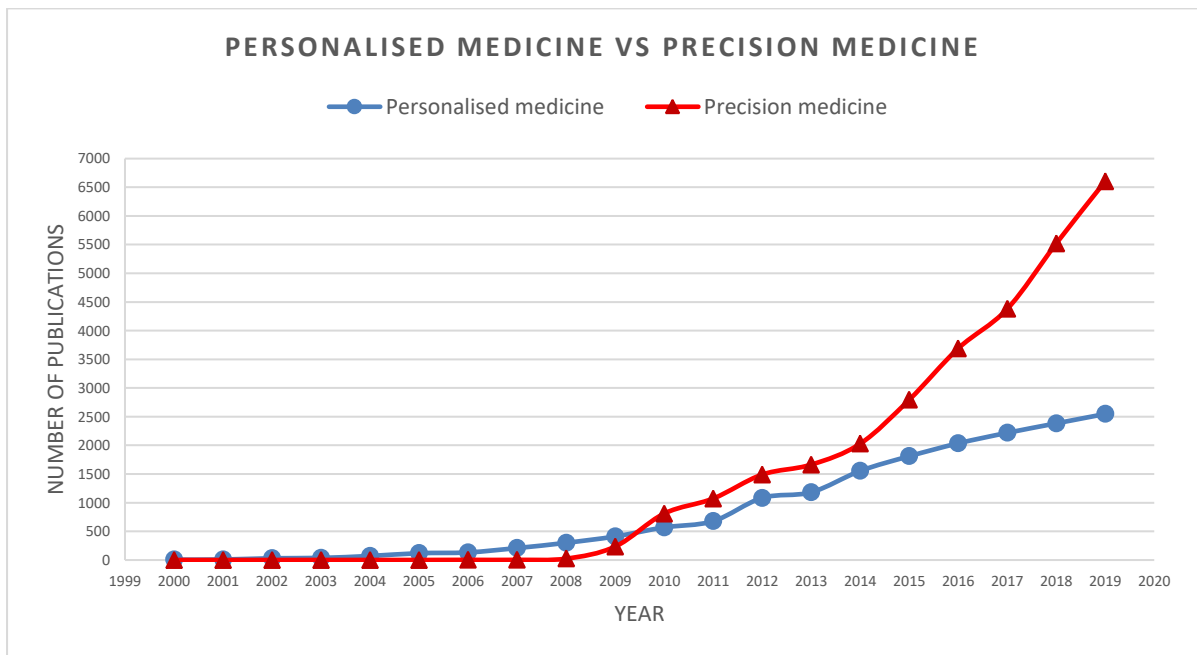
#### บทนำ

Precision medicine หรือการแพทย์แม่นยำเป็นศาสตร์แขนงใหม่ที่ต่อยอดเพิ่มเติมมาจากแนวความคิด personalised medicine โดยคำว่า personalised medicine ได้ถูกนำมาใช้และนำเสนออย่างเป็นทางการครั้งแรกต่อสาธารณชนจากบทความเรื่อง New Era of Personalised Medicine: Targeting Drugs for Each Unique Genetic Profile ซึ่งถูกตีพิมพ์ในปีค.ศ. 1999 ในช่วงแรกเริ่มนั้น กรอบแนวคิดของ personalised medicine มุ่งเน้นไปที่การให้การรักษาแก่ผู้ป่วยโดยคำนึงและตระหนักถึงความแตกต่างระหว่างบุคคลในหลากหลายมิติ อาทิเช่น พันธุกรรม (genetics) พันธุศาสตร์ด้านกระบวนการเหนือพันธุกรรม (epigenetic) อาการแสดงอื่นจากโรคร่วมและปัจจัยทางสภาพแวดล้อม เป็นต้น ซึ่ง personalised medicine ผนวกรวมทุกปัจจัยเพื่อถ่วงดุลแนวทางการรักษาที่เหมาะสมที่สุดต่อตัวบุคคลโดยเฉพาะ อย่างไรก็ตามในทางปฏิบัติ personalised medicine จะพิจารณาแผนการรักษาในแต่ละบุคคลโดยอ้างอิงข้อมูลเชิงประชากรที่มีลักษณะร่วมกันเพื่อวางแผนตัดสินใจการรักษาแล้วจึงปรับเปลี่ยนแผนการรักษาให้เหมาะสมต่ออาการของผู้ป่วยในแต่ละระยะการรักษา มิใช่การออกแบบจำเพาะต่อแต่ละปัจเจกบุคคลแต่อย่างใด

ดังนั้น ในระยะต่อมาก็เกิดคำนิยามใหม่ที่ชื่อว่า precision medicine ซึ่งมีคำจำกัดความว่า “การรักษาแบบแม่นยำจำเพาะเป็นแนวทางการรักษาที่คำนึงถึงความแตกต่างของปัจเจกบุคคลในมิติเชิงพันธุกรรม สภาพแวดล้อม และรูปแบบการใช้ชีวิตของแต่ละบุคคล” คำจำกัดความดังกล่าวไม่ได้แตกต่างกันมากนักเมื่อเทียบกับกรอบแนวคิดของ personalised medicine หากแต่คำว่า precision medicine ได้ถูกนำมาใช้มากขึ้นในช่วงระยะเวลา 20 ปีที่ผ่านมา (ระหว่างปี ค.ศ.1999 - 2019) (รูปที่ 1) เพื่อลดความสับสนด้านการแปลความเชิงไวยากรณ์และอิทธิพลจากผลการก่อตั้งโครงการ US Precision Medicine Initiative program ในปี 2006 ที่ประเทศสหรัฐอเมริกา ภายใต้การบริหารประเทศโดยประธานาธิบดีบารัค โอบามา ตัวอย่างของบริบทของ personalised/precision medicine ที่มีการประยุกต์ใช้จริง เช่น กรณีของการใช้ยา trastuzumab สำหรับรักษามะเร็งเต้านม หรือการใช้ยา ivacaftor และ lumacaftor เพื่อการรักษา cystic fibrosis

Trastuzumab มีข้อบ่งใช้เพื่อการรักษาผู้ป่วยมะเร็งเต้านมที่มีการแสดงออกของยีน human epidermal growth factor receptor subtype 2 (HER2 positive) โดยผลการวิจัยมีรายงานชัดเจนถึง

ประสิทธิภาพที่สูงกว่าอย่างมีนัยสำคัญทางสถิติในการรักษาผู้ป่วยมะเร็งเต้านมที่มี HER2 positive เมื่อเทียบกับผู้ป่วยที่ไม่มีการแสดงออกของยีน<sup>[1,2]</sup> หรืออีกกรณีหนึ่งที่มีความคล้ายคลึงกันคือ การเลือกใช้ยา ivacaftor ร่วมกับยา lumacaftor ในการรักษาผู้ป่วยที่มีภาวะ cystic fibrosis ที่มีการแสดงออกของ cystic fibrosis transmembrane conductance regulator (CFTR) proteins แบบ F508del missense mutation ซึ่งพบได้มากถึง 88% ในผู้ป่วย cystic fibrosis<sup>[3]</sup> โดยพบว่าในผู้ป่วยที่มี mutation ลักษณะดังกล่าว การใช้ยาสองชนิดร่วมกันให้ประสิทธิผลการรักษาสูงสุดเมื่อเทียบกับการให้การรักษาด้วยยาตัวใดตัวหนึ่งเป็นยาเดี่ยว<sup>[4]</sup> ตัวอย่างข้างต้นเป็นการนำเสนอความสำคัญของการผนวกความรู้ด้านเภสัชพันธุศาสตร์ (pharmacogenomics) ประกอบการตัดสินใจเลือกใช้ยา (medical decision-making) ซึ่งการตัดสินใจดังกล่าวตั้งอยู่บนพื้นฐานของการใช้ตัวชี้วัดทางชีวภาพ (biomarkers) ต่อทิศทางการรักษาผู้ป่วย ซึ่งถือว่าอยู่ในบริบทของการแพทย์แม่นยำจำเพาะ



รูปที่ 1. แสดงผลการค้นหาบนฐานข้อมูล PubMed ด้วยการใช้คำสืบค้น “Personalised medicine” และ “Precision medicine” ช่วงระยะเวลาระหว่างปีค.ศ. 1999 - 2019

### ความก้าวหน้าของวิทยาการด้านข้อมูลและปัญญาประดิษฐ์ต่อพัฒนาการของ precision medicine

วิทยาการที่ถูกกล่าวถึงอยู่เสมอเกี่ยวกับ precision medicine คือ วิทยาการด้านเภสัชพันธุศาสตร์ (pharmacogenomics) ซึ่งมีการศึกษามายาวนานและเป็นหนึ่งในแขนงวิชาที่ประสบผลสำเร็จอย่างสูง

เนื่องจากการนำความรู้ที่สังเคราะห์ได้มาประยุกต์ใช้จริงในหลากหลายสถานการณ์ เช่น ศึกษาผลของ genetic polymorphism ของ *CYP2C19* ต่อการออกฤทธิ์ของยา clopidogrel หรือการมีอัตราการแปรสภาพยาที่ลดลง (poor metaboliser) ของยา warfarin เมื่อเทียบกับผู้ป่วยทั่วไปอื่นเนื่องจากการแสดงออกของ gene *CYP2C9\*2* <sup>[5,6]</sup> อย่างไรก็ตาม pharmacogenomics ส่วนมากมักใช้การทดลองในห้องปฏิบัติการ เช่น การวิจัยในเนื้อเยื่อ (tissue cultures) หรือเก็บตัวอย่างจากประชากรจริง (*ex vivo*) แล้วนำมาทำการทดลองเชิงชีววิทยาระดับโมเลกุล เช่น Western blot, Real-time PCR, Next generation sequencing หรือการวิจัยเชิงโอมิกส์ (omics) เช่น metabolomics, proteomics อีกหลายขั้นตอนจึงจะสามารถสรุปผลประกอบการนำไปประยุกต์ใช้จริงได้<sup>[7]</sup> แนวทางการวิจัยดังกล่าวสำหรับประเทศกำลังพัฒนาแล้วถือว่าเป็นประเด็นที่มีความท้าทายเป็นอย่างสูง<sup>[8-10]</sup> เนื่องจากการทดลองทางห้องปฏิบัติการขั้นสูงเหล่านั้นจำเป็นต้องใช้ทรัพยากรทางการเงินและทรัพยากรทางบุคคลในปริมาณที่สูง อีกนัยหนึ่ง การทดลองบางประเภทอาจต้องใช้ระยะเวลาอย่างยาวนานจึงจะสามารถบรรลุผลสัมฤทธิ์และพิสูจน์สมมติฐานที่ตั้งไว้ได้ งานวิจัยยังซับซ้อนก็ยิ่งส่งผลให้เกิดความล้มเหลวได้อันเนื่องมาจากปัจจัยของ random error, human error และ systematic error ดังนั้นโอกาสล้มเหลวในการทดลองก็จัดเป็นความเสี่ยงหนึ่งที่นักวิจัยมักเผชิญหน้าตลอดเวลา<sup>[11,12]</sup>

จากข้อจำกัดดังกล่าวจึงทำให้เกิดแนวความคิดการใช้ปัญญาประดิษฐ์และแบบจำลองทางคอมพิวเตอร์เพื่อช่วยสนับสนุนการวิจัย<sup>[13-15]</sup> แนวคิดดังกล่าวเกิดขึ้นพร้อม ๆ กับความก้าวหน้าทางวิทยาการคอมพิวเตอร์ในโลกยุคปัจจุบัน ซึ่งผลของความก้าวหน้าดังกล่าวได้มีอิทธิพลต่อวิทยาการทางการแพทย์ทั้งในเชิงคลินิกและวิทยาศาสตร์การแพทย์พื้นฐาน เช่นเดียวกัน วิทยาการคอมพิวเตอร์ได้ถูกนำมาประยุกต์เข้ากับเภสัชวิทยาเชิงคลินิก (clinical pharmacology) เช่น การสร้างแบบจำลองการกระจายของยาในร่างกาย (physiologically-based pharmacokinetics modelling, PBPK) การทำ simulation ประกอบการตัดสินใจการเลือกขนาดของยาเริ่มต้นในผู้ป่วยกลุ่มพิเศษ และการใช้ระบาดวิทยาทางคลินิก (clinical epidemiology) ในการวิเคราะห์ข้อมูลเพื่อระบุตัวบ่งชี้ (predictors) ที่ใช้พยากรณ์ผลการรักษาด้วยยา ในบทความนี้จะขอยกตัวอย่างในส่วนของระบาดวิทยาคลินิก และ PBPK ที่มีส่วนส่งเสริมความก้าวหน้าของแนวทาง precision medicine

### ระบาดวิทยาคลินิก (Clinical epidemiology)

ระบาดวิทยาคลินิกถูกประยุกต์เข้ากับแนวคิดด้าน precision medicine มากขึ้นในระยะเวลาที่ผ่านมา เนื่องมาจากแก่นหลักของระบาดวิทยาคลินิกคือ การปรับปรุงและเปรียบเทียบผลการรักษาด้วยวิธีการใหม่เทียบกับวิธีการที่มีก่อนหน้า ซึ่งถือว่าเป็นไปตามครรลองของ precision medicine เป็นอย่างดี<sup>[16-18]</sup> ศาสตร์ด้านนี้มุ่งเน้นการใช้ความรู้ทางคณิตศาสตร์และสถิติ เช่น linear regression, logistic regression และ Bayesian interference เพื่อทดสอบสมมติฐานของงานวิจัย ซึ่งวิทยาการข้อมูล (data science) มีส่วนช่วย

อย่างมากต่อการคำนวณวิเคราะห์และสร้างแบบจำลองทางคณิตศาสตร์จากข้อมูลขนาดใหญ่ (big data) ซึ่งนำไปสู่การต่อยอดทาง precision medicine ได้<sup>[18,19]</sup>

### กรณีศึกษาที่ 1

กรณีศึกษาที่ 1 จะเป็นการนำเสนอการใช้ R programming language (<https://www.r-project.org/>) เพื่อวิเคราะห์สมมติฐานต่าง ๆ เกี่ยวกับผลการรักษาผู้ป่วยมะเร็งระดับด้วยยา sorafenib ซึ่งเป็นยาในกลุ่ม tyrosine kinase inhibitors (TKIs) โดยข้อมูลที่ใช้วิเคราะห์เป็นข้อมูลจริงจากการทดลอง phase III clinical trial NCT00699374<sup>[20]</sup> ผ่านทาง data sharing platform: Project Data Sphere<sup>®</sup> (<https://www.projectdatasphere.org/>) ข้อมูลที่ได้รับจากฐานข้อมูลมีลักษณะซับซ้อน (ตารางที่ 1) ประกอบด้วยไฟล์ย่อยที่บรรจุข้อมูลต่างชนิดกัน เช่น ข้อมูลพื้นฐานของผู้ป่วยแต่ละรายซึ่งไม่ระบุตัวตน (ถูกเข้ารหัสด้วย enrolment ID) ข้อมูลอาการข้างเคียง ข้อมูลยาที่มีการใช้ควบคู่ระหว่างเข้าร่วม clinical trial เป็นต้น หลังจากได้รับข้อมูลชุดดังกล่าวจากฐานข้อมูล การแปรสภาพให้อยู่ในรูปแบบที่เหมาะสมต่อการวิเคราะห์ (data wrangling) (รูปที่ 2) ถูกดำเนินการบนโปรแกรม RStudio ซึ่งทำงานร่วมกับ โปรแกรม R (จัดเป็นภาษาทางคอมพิวเตอร์ชนิดหนึ่ง) ในระหว่างการเขียนคำสั่ง (coding) เพื่อแปรสภาพข้อมูลนั้นก็ สามารถเลือกที่จะผนวกรวมข้อมูลที่จำเป็นเข้าเป็นไฟล์เดียวกัน (merging) ซึ่งทำให้ง่ายต่อการวิเคราะห์ความสัมพันธ์ของตัวแปรแต่ละตัวต่อสมมติฐานที่ตั้งไว้ โดยสมมติฐานที่ถูกตั้งก่อนการวิเคราะห์ข้อมูลชุดนี้คือ

1. การปรับเปลี่ยนขนาดยาตั้งแต่เริ่มต้นอาจส่งผลไม่มากนักน้อยต่อประสิทธิภาพการรักษาและตัวแปรใดบ้างที่ช่วยบ่งชี้ถึงแนวโน้มที่จะนำไปสู่การปรับเปลี่ยนขนาดยา
2. การใช้ Proton pump inhibitors (PPIs) มีผลลดการดูดซึมของยา sorafenib เนื่องจากเคยมีรายงานอันตรกิริยาของ PPIs ต่อการลดการดูดซึมของยาในกลุ่ม TKIs ตัวอื่นๆ

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC
1	STUDYID	SUBJID	RANDDT	TRTSTDT	TRTENDT	RFSTDT	RFENDT	ITTF	SAFFL	ARM	AGE	AGE.decat	SEX	WT	HT	RACE	RACE.ASI	RACE.WHI	RANDDY	RFSTDY	RFENDY	TRTENDY	STUDYENF	STUDYENF	DXDY	DXYR	ECOGPS	SRVDY	SRVCFL
2	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	62	6.2 Female	42.9	161.8	White	Non-Asiar	White	1	1	139	139	140	OBJECTIV	140	0.383299	1	196	Event
3	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	83	8.3 Female	74.7	175.4	White	Non-Asiar	White	0	1	249	249	251	OBJECTIV	550	1.505818	0	447	Event
4	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	67	6.7 Female	48.5	156	White	Non-Asiar	White	1	1	357	357	358	OBJECTIV	20	0.054757	0	928	Event
5	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	68	6.8 Male	97.5	177	White	Non-Asiar	White	1	1	59	59	59	PROTOCO	NA	NA	0	1011	Censorec
6	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	73	7.3 Male	56.5	167	Asian	Asian	Non-Whit	1	1	351	351	363	OBJECTIV	1250	3.422313	0	615	Event
7	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	60	6 Male	67.3	165.9	Asian	Asian	Non-Whit	-2	1	29	29	44	OBJECTIV	935	2.55989	0	61	Event
8	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	73	7.3 Male	60.2	160	Asian	Asian	Non-Whit	-1	1	178	178	180	ADVERSE	1737	4.755647	0	273	Event
9	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	82	8.2 Male	50.5	162.4	Asian	Asian	Non-Whit	0	1	11	11	29	ADVERSE	1639	4.487337	0	668	Censorec
0	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	52	5.2 Male	55	161.9	Asian	Asian	Non-Whit	0	1	49	49	49	OBJECTIV	18	0.049281	0	339	Event
1	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	69	6.9 Male	56.1	155	Asian	Asian	Non-Whit	-2	1	29	29	44	OBJECTIV	9	0.024641	0	412	Event
2	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	67	6.7 Male	61	156.8	Asian	Asian	Non-Whit	-5	1	113	113	128	OBJECTIV	1671	4.574949	0	625	Censorec
3	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	56	5.6 Male	68.4	172.7	Asian	Asian	Non-Whit	-4	1	45	45	45	OBJECTIV	10	0.027379	1	69	Event
4	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	60	6 Female	57.9	154.9	Asian	Asian	Non-Whit	0	1	50	50	57	OBJECTIV	165	0.451745	1	166	Event
5	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	69	6.9 Male	64.2	160	Asian	Asian	Non-Whit	0	1	123	123	124	OBJECTIV	238	0.651608	0	173	Event
6	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	73	7.3 Female	41.9	147.3	Asian	Asian	Non-Whit	0	1	15	15	36	NO LONG	12	0.032854	0	354	Censorec
7	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	50	5 Male	51	164.5	Asian	Asian	Non-Whit	-3	1	29	29	53	OBJECTIV	237	0.648871	0	82	Event
8	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	54	5.4 Male	77.1	176.1	Asian	Asian	Non-Whit	-1	1	10	10	30	ADVERSE	1037	2.839151	0	765	Censorec
9	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	69	6.9 Male	58	162	Asian	Asian	Non-Whit	1	1	24	24	50	OBJECTIV	11	0.030116	0	684	Censorec
0	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	58	5.8 Male	60.1	163.8	Asian	Asian	Non-Whit	0	1	253	253	253	OBJECTIV	63	0.172485	0	733	Censorec
1	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	72	7.2 Female	43.6	149.6	Asian	Asian	Non-Whit	0	1	61	61	61	OBJECTIV	0	0	0	176	Event
2	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	64	6.4 Male	67.1	171.7	Asian	Asian	Non-Whit	0	1	47	47	54	OBJECTIV	364	0.996578	0	150	Event
3	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	54	5.4 Male	47.8	167.4	Asian	Asian	Non-Whit	0	1	83	83	83	OBJECTIV	1783	4.881588	0	184	Event
4	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	78	7.8 Male	61.7	161.2	Asian	Asian	Non-Whit	-2	1	85	85	85	OBJECTIV	18	0.049281	0	137	Event
5	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	68	6.8 Female	67	145.9	Asian	Asian	Non-Whit	0	1	168	168	169	OBJECTIV	1526	4.17796	0	593	Event
6	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	74	7.4 Male	44.4	155.4	Asian	Asian	Non-Whit	0	1	702	702	703	STUDY TEF	1375	3.764545	0	731	Censorec
7	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	69	6.9 Male	64.4	170.8	Asian	Asian	Non-Whit	0	1	127	127	137	OBJECTIV	10	0.027379	0	264	Event
8	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	66	6.6 Male	58	161.5	Asian	Asian	Non-Whit	-2	1	128	128	141	OBJECTIV	1987	5.44011	0	479	Event
9	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	71	7.1 Male	54	175.3	Asian	Asian	Non-Whit	-17	1	81	81	82	OBJECTIV	1358	3.718001	1	285	Event
0	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	59	5.9 Male	52.9	161.4	Asian	Asian	Non-Whit	-2	1	86	86	86	OBJECTIV	75	0.205339	1	397	Censorec
1	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	46	4.6 Male	65.1	178	Asian	Asian	Non-Whit	0	1	44	44	45	OBJECTIV	8	0.021903	0	469	Censorec
2	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	71	7.1 Male	67.4	172.6	Asian	Asian	Non-Whit	0	1	112	112	113	ADVERSE	510	1.396304	0	473	Censorec
3	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	73	7.3 Male	49.6	161.4	Asian	Asian	Non-Whit	1	1	80	80	81	OBJECTIV	11	0.030116	0	589	Event
4	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	70	7 Female	54	154.7	Asian	Asian	Non-Whit	0	1	62	62	86	ADVERSE	1	0.002738	0	589	Censorec
5	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	56	5.6 Male	60.5	165.7	Asian	Asian	Non-Whit	-1	1	277	277	278	OBJECTIV	255	0.698152	0	572	Event
6	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	48	4.8 Female	51.3	157	Asian	Asian	Non-Whit	1	1	138	138	211	SUBJECT	1689	4.62423	0	211	Event
7	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	56	5.6 Female	46.6	164.4	Asian	Asian	Non-Whit	1	1	81	81	82	OBJECTIV	83	0.227242	0	241	Event
8	A6181170	A6181170	#####	#####	#####	#####	#####	#####	Y	Y	Sorafenib	78	7.8 Male	59.4	155.6	Asian	Asian	Non-Whit	1	1	11	11	15	ADVERSE	930	2.546201	0	35	Event

ตารางที่ 1. ข้อมูลดิบที่ได้จาก data sharing platform (Project datasphere)

The image shows the RStudio interface with four main components labeled A, B, C, and D:

- A:** The source editor showing R code for calculating Gini coefficients, ranking features, and fitting a logistic model.
- B:** The Environment pane showing loaded objects: data (105471 obs. of 771 variables), ds (105471 obs. of 771 variables), dt (100 obs. of 3 variables), sampled (5000 obs. of 3 variables), and subset\_loss (9783 obs. of 771 variables).
- C:** The Files pane showing a histogram plot titled "Histogram of subset\_loss\$loss".
- D:** The Console pane showing the execution of R commands and their output, including a summary of loss and histograms for the data and subset\_loss.

```
135 Gini_value = Gini(predicted_values, ds$default);
136 performance_metric_gini = rbind(performance_metric_gini, cbind(feature, Gini_value));, error=function(e
137 ){}
138 }
139 }
140 performance_metric_gini$Gini_value = as.numeric(as.character(performance_metric_gini$Gini_value))
141
142 # rank the Features by value of Gini coefficient
143 Ranked_Features = performance_metric_gini[order(-performance_metric_gini$Gini_value),]
144 print("Top 5 Features by Gini coefficient\n")
145 head(Ranked_Features)
146
147 # The ranking methods tells us that the top six features by their individual predicted
148 # power are f766, f404, f629, f630, f281, and f322. The top feature in the Gini coefficient is
149 # 0.268 (or 26.8%)
150
151 # Now using the set of top five features, let's create a logistical model and
152 # see the same performance metric
153
154 # Create logistic model
155 glm_model = glm(default ~ f766 +f404 +f629 +f630 +f281+ f322, data = ds, family = binomial(link = 'logit'));
156 predicted_values = predict(glm_model, newdata = ds, type = 'response');
157 Gini_value = Gini(predicted_values, ds$default);
158 summary(glm_model)
159 Gini_value
160 # The model result shows that four features (f766, f404, f630, and f281) are not significant
161 # This gives us an indication that the features themselves are highly correlated and hence are not adding
162 # value by being in the model.
```

```
> summary(ds$loss)
  Min. 1st Qu.  Median    Mean 3rd Qu.   Max.
0.0000 0.0000 0.0000 0.7996 0.0000 100.0000
> hist(data$loss,
+       xlab = "Loss",
+       border = 'blue',
+       col = 'green',
+       las = 0,
+       breaks = 100,
+       probability = TRUE)
> # subset data into non-loss and loss (ie, loss > 0)
> subset_loss = subset(ds, loss != 0)
> hist(subset_loss$loss, xlab = "Loss",
+       border = 'blue',
+       col = 'green',
+       las = 0,
+       breaks = 100,
+       probability = TRUE)
```

รูปที่ 2. ภาพรวมการเขียนคำสั่งด้วย R บน RStudio environment. A: source, B: environment, history and others,

C: files, plot and installed packages, D: console

## ตัวอย่างการเขียนคำสั่งบน R

### 1.1. นำเข้าข้อมูลหลัก (Data preparation)

```
set.default.data.dirpath("../Sorafenib HCC/data")  
dataset = get.table("SL") #loads the cleaned table containing subject level variables - i.e. variables  
that have only one value per subject e.g. age, sex, race
```

### 1.2. นำเข้าตารางข้อมูลรวมที่จำเป็น (Load tables)

```
tblAE = get.table("AE") # นำเข้าตารางรายงานอาการข้างเคียงของผู้ป่วยแต่ละราย  
tblCM = get.table("CM") # นำเข้าประวัติยาที่ใช้ร่วมระหว่างการศึกษ  
tblLB = get.table("LB") # นำเข้าผลตรวจทางห้องปฏิบัติการ  
tblMH = get.table("MH") # นำเข้าประวัติการใช้ยาเดิมของผู้ป่วยก่อนเข้าร่วมการศึกษา
```

### 1.3. ตรวจสอบภาพรวมของข้อมูล (General summarisation of dataset)

```
summary(dataset)  
describe(dataset)  
summary(dataset$DXYR) # summarise only Years Since Diagnosis (DXYR)  
describe(dataset$DXYR)
```

### 1.4. สรุปข้อมูลพื้นฐานต่างๆแล้วจัดเก็บเป็นข้อมูลแยก

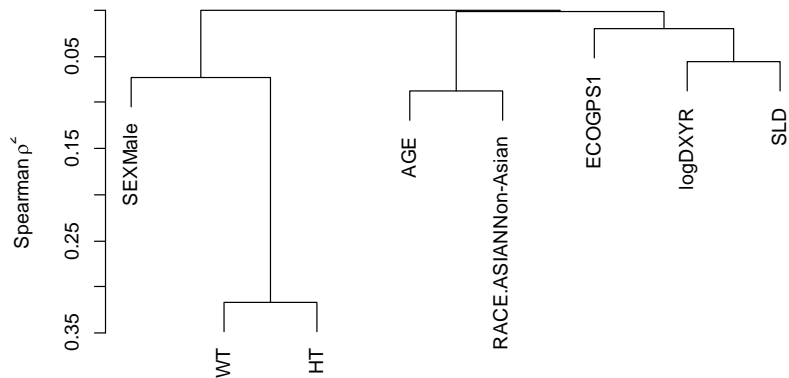
```
descriptive_report(data = dataset,var = c("SEX", "AGE", "WT", "HT", "RACE.ASIAN", "SLD", "DXYR",  
"ECOGPS", "WBC.HIFL.BL"),filename = "descriptive statistics table.html")
```

### 1.5. สร้างแผนภาพต้นไม้เพื่อประเมินทิศทางการความสัมพันธ์ของตัวแปรที่สนใจด้วยการใช้สหสัมพันธ์เชิงอันดับสเปียร์แมน (Spearman's correlation)



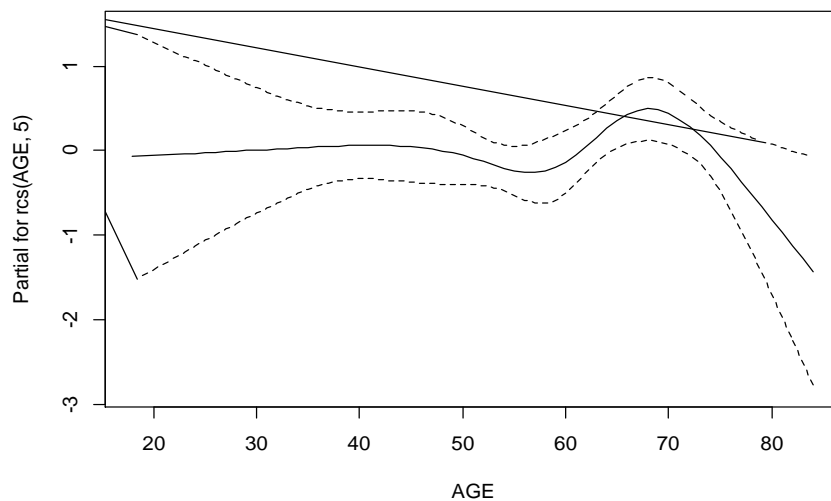
```
vc = varclus(all.vars.formula, sim='spearman', data=dataset)
```

```
plot(vc)
```



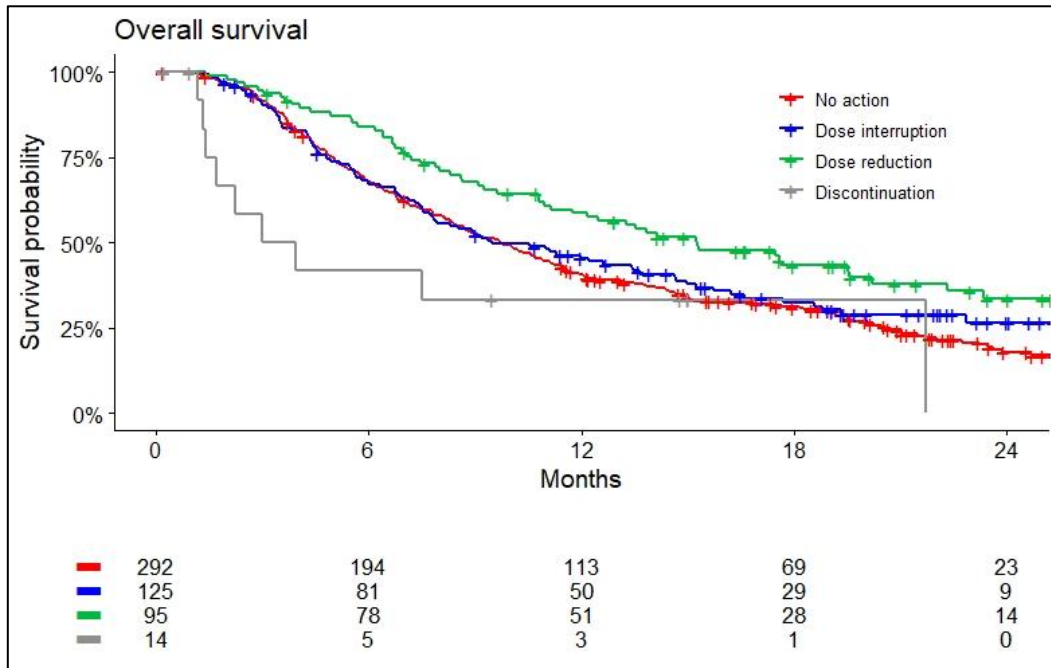
1.6. ตรวจสอบรูปแบบความสัมพันธ์ของข้อมูลพื้นฐาน (อายุ) ต่อการเกิดอาการข้างเคียงขั้นรุนแรง (hand-foot syndrome grade 3 or greater) จากการใช้ยา sorafenib

```
test.continuous(lr(l(HANDSYND.MAXGR >= 3) ~ 1, data=dataset), 'AGE') # กราฟที่ได้จากการทดสอบ (กราฟที่ A) บ่งชี้ถึงความสัมพันธ์ไม่เป็นเส้นตรงระหว่างตัวแปรที่ทำการทดสอบ (non-linear relationship)
```

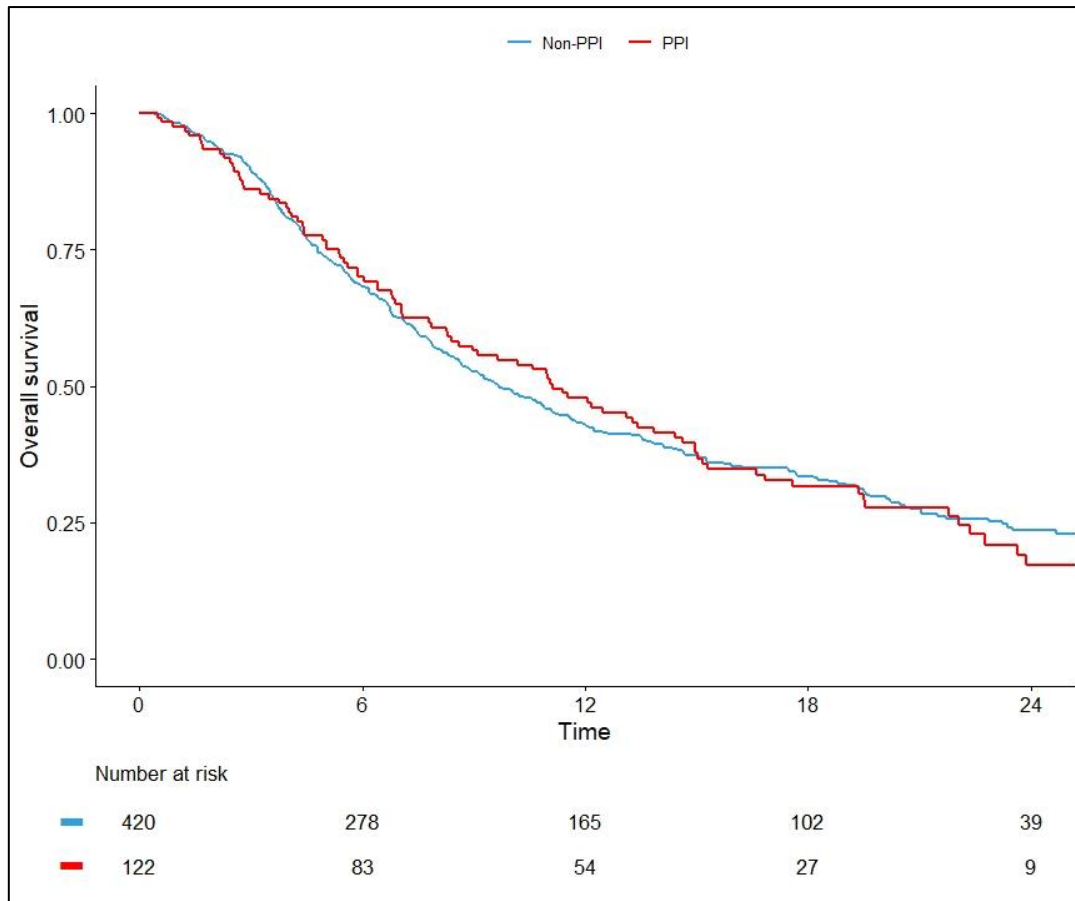


ซึ่งจากผลการวิเคราะห์โดยสรุปได้ผลทดสอบสมมติฐานคือ 1. การปรับขนาดยา ไม่มีผลกระทบต่อผลการรักษาในภาพรวม และสามารถใช้อัตราพื้นฐานของผู้ป่วย เช่น เชื้อชาติ เพศและอายุ ในการทำนายแนวโน้มการเกิดอาการข้างเคียงซึ่งนำไปสู่การปรับเปลี่ยนขนาดของยา sorafenib ได้ (รูปที่ 3)<sup>[21]</sup> 2. การใช้

PPIs ควบคู่ไปกับ sorafenib ไม่ส่งผลต่อประสิทธิภาพของการรักษาแต่อย่างใด ซึ่งส่งผลดีต่อความเชื่อมั่นในการใช้ยาทั้งตัวผู้ป่วยและแพทย์ผู้รักษา ในกรณีที่ผู้ป่วยจำเป็นที่จะต้องได้รับยา PPIs เพื่อรักษาอาการทางระบบทางเดินอาหาร (รูปที่ 4)<sup>[22]</sup>



รูปที่ 3. Kaplan-Meier แสดง overall survival ในผู้ป่วยแยกตามลักษณะของการปรับยา sorafenib ในผู้ป่วยมะเร็งตับ (HR [95% CI]; dose interruption = 0.9 [0.7-1.2]; dose reduction = 0.6 [0.5-0.9]; discontinuation = 1.7 [0.9-3.4]; P = 0.005)<sup>[21]</sup>



รูปที่ 4. Kaplan-Meier แสดง overall survival ระหว่างกลุ่มที่ใช้ PPIs และ ไม่ใช้ PPIs ร่วมกับยา sorafenib HR [95% CI] 1.01 [0.80–1.28],  $P = 0.93$ ) ในผู้ป่วยมะเร็งตับ<sup>[22]</sup>

จะเห็นได้ว่าการใช้วิทยาการข้อมูลมีส่วนช่วยอย่างมากต่อการวิเคราะห์ข้อมูลประเภท big data ที่เกี่ยวกับการใช้ยา ข้อมูลดิบจาก clinical trial จัดเป็นข้อมูลที่มีคุณภาพสูง เนื่องจากการทำการทดลองยาในมนุษย์ แต่ละโครงการวิจัยจำเป็นที่จะต้องผ่านมาตรฐานการรับรองคุณภาพในหลายภาคส่วนจึงจะได้รับอนุมัติให้ดำเนินการได้ การนำข้อมูลประเภทนี้มาวิเคราะห์ซ้ำ (secondary analysis) โดยมุ่งเป้าทดสอบสมมติฐานอื่น ๆ นอกเหนือไปจากผลการทดลองหลัก ส่งผลให้เกิดองค์ความรู้ใหม่ที่สามารถนำไปใช้ได้จริงและมีประโยชน์ต่อบุคลากรทางสุขภาพที่เกี่ยวข้อง

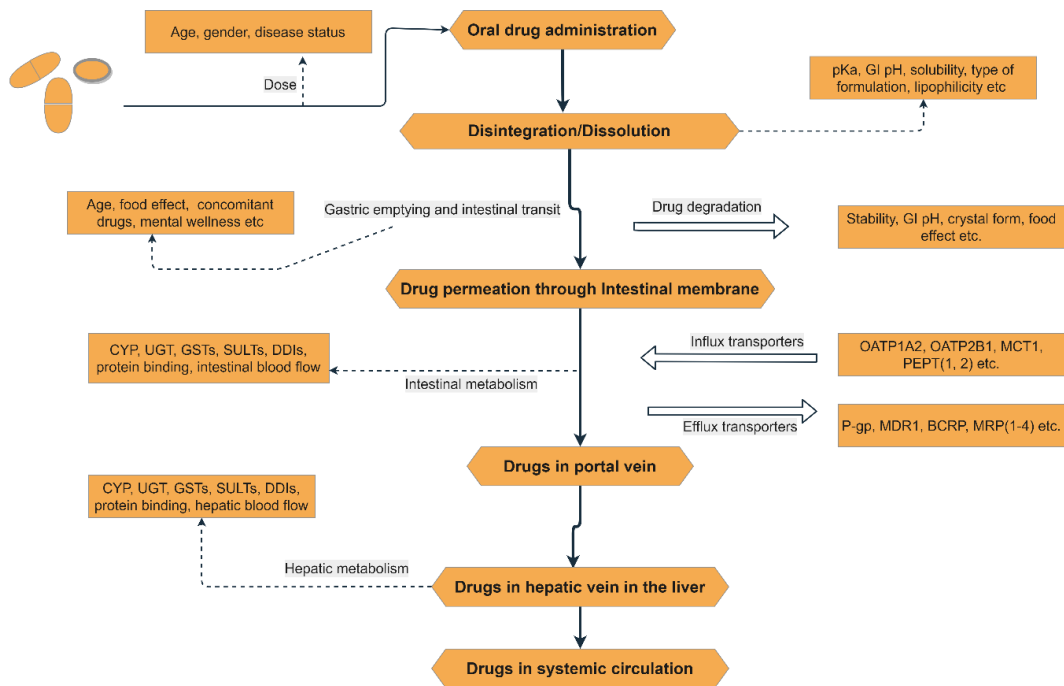
### Physiologically-based pharmacokinetics modelling and simulation (PBPK)

PBPK หรือ แบบจำลองเภสัชจลนศาสตร์บนฐานสรีรวิทยา เป็นศาสตร์อีกแขนงหนึ่งที่ใช้ประโยชน์จากความก้าวหน้าทางวิทยาการคอมพิวเตอร์ในการสร้างแบบจำลองทางสรีรวิทยาเพื่ออธิบายผลทางเภสัชจลนศาสตร์หลังจากบริหารยาเข้าสู่ร่างกาย แบบจำลอง PBPK จะแบ่งอวัยวะของร่างกายออกเป็น compartments หลาย ๆ ส่วน ซึ่งแต่ละส่วนก็จะมีลักษณะเฉพาะ (characteristic) เช่น อัตราการดูดซึม อัตราการกำจัดออกของยา หรือ อัตราการไหลเวียนของเลือดแตกต่างกัน ในปัจจุบัน PBPK มักจะถูกสร้างออกมาเป็นโปรแกรมสำเร็จรูปพร้อมต่อการใช้งาน อาทิเช่น GastroPlus<sup>®</sup>, Simcyp<sup>®</sup> และ PK-sim<sup>®</sup> โดย Simcyp<sup>®</sup> เป็นหนึ่งในโปรแกรมที่ได้รับการยอมรับจากสำนักงานคณะกรรมการอาหารและยา ประเทศสหรัฐอเมริกา (US FDA) ให้สามารถจำลองข้อมูลประกอบการขออนุมัติการขึ้นทะเบียนยาใหม่ ในกรณีของ Simcyp<sup>®</sup> ตัวโปรแกรมจะคำนึงถึงส่วนหลักสองส่วนใหญ่ ได้แก่ ข้อมูลจำเพาะของยา (เช่น dosage form, dissolution, permeation, protein binding, bioavailability) และ ข้อมูลของประชากรจำลอง (เช่น อายุ, เชื้อชาติ, การแสดงออกของ drug-metabolising enzymes, การทำงานของไต) แล้วจึงทำการจำลองการบริหารยาสมมติ (virtual clinical trials) ด้วยแบบจำลองยา (compound models) และ ผู้ป่วยจำลอง (virtual populations) ที่สร้างขึ้น

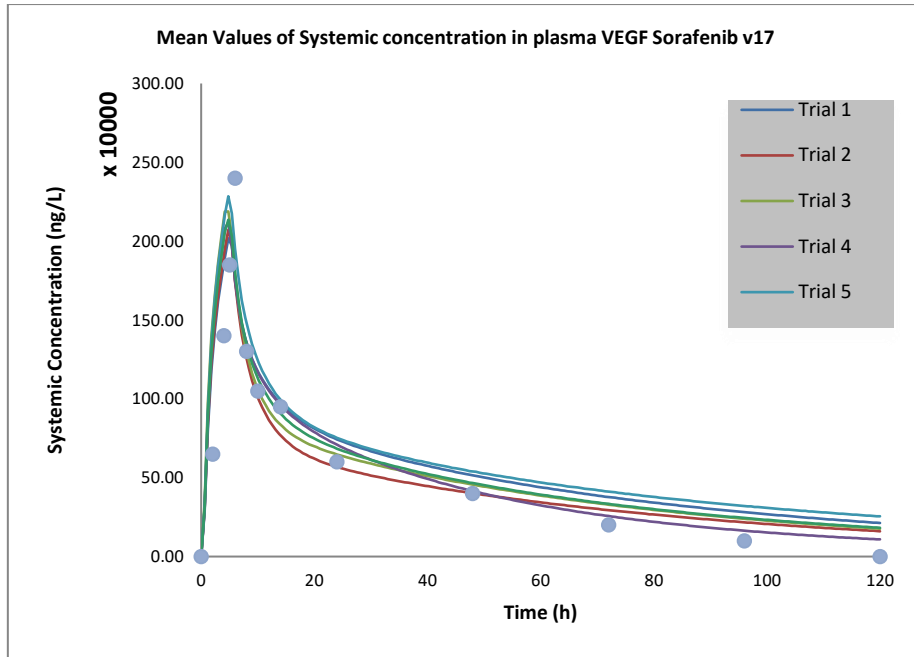
#### กรณีศึกษาที่ 2

กรณีศึกษาที่ 2 จะเป็นการใช้ Simcyp<sup>®</sup> เพื่อจำลองลักษณะการตอบสนองของยา sorafenib โดยมุ่งแยกแยะหาปัจจัยทางสรีรวิทยาหรือลักษณะของยาที่ส่งผลกระทบต่อความแปรปรวนระหว่างบุคคลของยา sorafenib<sup>[23]</sup> ซึ่งตามรายงานการใช้ยา sorafenib นั้น พบการแปรปรวนของการตอบสนองของยาที่ค่อนข้างสูง ในการศึกษานี้ได้ทำการสร้าง sorafenib compound model โดยอ้างอิงจากฐานข้อมูลของ US FDA และงานวิจัยใน phase I และ phase II ของยา sorafenib หลังจากนั้นจึงเข้าสู่ขั้นตอนการ validate คุณลักษณะทางเภสัชจลนศาสตร์ของตัวยาจำลอง โดยทำการ simulation เบื้องต้นเพื่อตรวจสอบพฤติกรรมของยา (รูปที่ 6) ในการจำลองข้อมูลส่วนนี้ตัวโปรแกรมจะมีตัวเลือกสำหรับรูปแบบการจำลอง (sub-model)

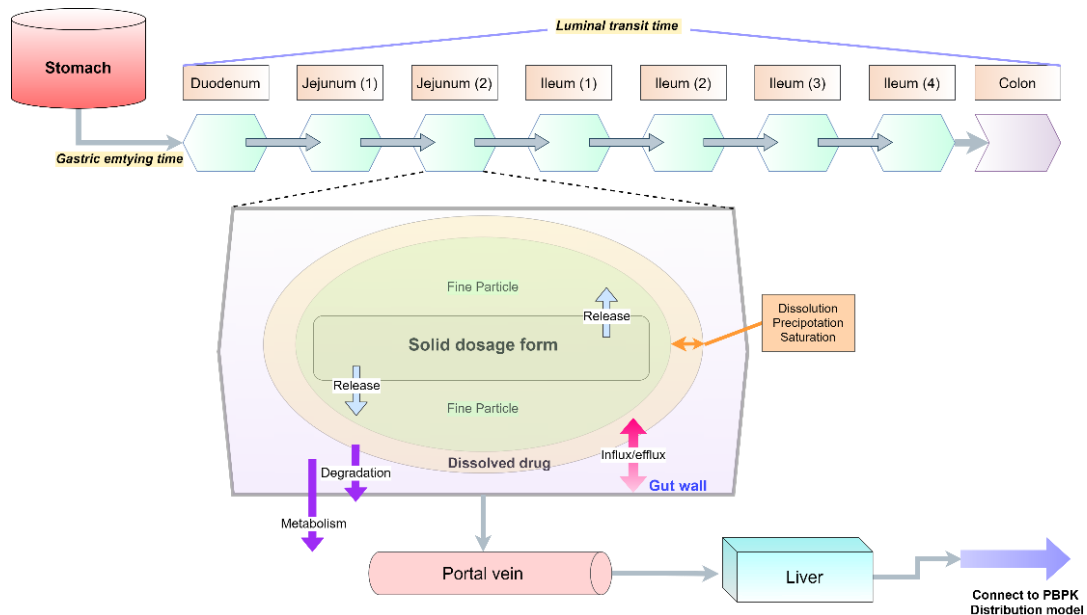
เช่น first-order kinetics, enhanced Compartmental Absorption and Transit (CAT), advanced compartmental absorption and transit (ACAT) model และ the advanced dissolution, absorption and metabolism (ADAM) model (รูปที่ 7) ซึ่ง ADAM sub-model เป็นรูปแบบจำลองที่เหมาะสมที่สุดในการจำแนกหาปัจจัยที่ส่งผลต่อความแปรปรวนการตอบสนองของยา เนื่องจาก ADAM sub-model จะคำนึงถึงความสัมพันธ์ของแต่ละอวัยวะที่มีอิทธิพลต่อกระบวนการ metabolism ของยา



รูปที่ 5. อันตรกิริยาและปัจจัยต่างๆของร่างกายที่มีอิทธิพลต่อการบริหารยาทางปาก ดังจะเห็นได้ว่าในแต่ละขั้นตอนต่างมีปัจจัยที่แตกต่างกันซึ่งส่งผลกระทบต่อเภสัชจลนศาสตร์และเภสัชพลศาสตร์ของยา



รูปที่ 6. ผลการจำลองพฤติกรรมทางเภสัชจลนศาสตร์ของ sorafenib compound model โดยใช้ Simcyp® software ซึ่งสามารถกำหนดลักษณะของผู้ป่วยจำลองไปจนถึงจำนวนครั้งของการจำลองข้อมูล ทั้งนี้เพื่อใช้ในการตรวจสอบพฤติกรรมทางเภสัชจลนศาสตร์พื้นฐานของ compound model ที่ถูกสร้างขึ้น



รูปที่ 7. โครงสร้างของ ADAM sub-model ซึ่งมีการแบ่งระบบทางเดินอาหารออกเป็น 9 compartments (stomach segment + 7 small intestine segments + colon segment) การดูดซึมในแต่ละส่วนจะ

ขึ้นอยู่กับลักษณะการปลดปล่อยของยาตามแต่ละ dosage form พฤติกรรมการละลายของยาไปจนถึงปัจจัยย่อยอื่นๆซึ่งผลรวมทั้งหมดในแต่ละ compartment ต่างมีอิทธิพลต่อ compartment ที่อยู่ในลำดับถัดไป หลังจากตรวจสอบพฤติกรรมทางเภสัชจลนศาสตร์ของ sorafenib compound model ขั้นตอนถัดไปจึงเป็นการจำลองในสถานการณ์ที่เสมือนจริงมากขึ้น โดยเพิ่มขนาดของประชากร (n = 1,000) และกำหนดรูปแบบการบริหารยา (dosing regimen) เสมือนกำลังทำโครงการ clinical trial ข้อมูลส่วนนี้จะถูกนำไปวิเคราะห์แยกแยะหาความสัมพันธ์ของปัจจัยทางสรีรวิทยาหรือลักษณะของยาที่ส่งผลกระทบต่อ การตอบสนองของยา ด้วยการทำ univariate และ multivariable linear regression model

ตารางที่ 2 : แสดงความสัมพันธ์ของปัจจัย (predictors) ทางสรีรวิทยาที่ส่งผลกระทบต่อ การตอบสนองของยา ด้วยกระบวนการ stepwise forward inclusion และรายงานประสิทธิภาพการพยากรณ์ของแต่ละ multivariable linear regression model

Model	R <sup>2</sup>	Std. Error of the Estimate	R <sup>2</sup> Change
a	0.631	0.24141	0.631
b	0.781	0.18614	0.150
c	0.868	0.14458	0.087
d	0.873	0.14156	0.006
e	0.883	0.13619	0.010
f	0.883	0.13595	0.001

Model predictors (a) hepatic CYP3A4 abundance; (b) hepatic CYP3A4 abundance, albumin concentration; (c) hepatic CYP3A4 abundance, albumin concentration, BMI; (d) hepatic CYP3A4 abundance, albumin concentration, BMI, body surface area; (e) hepatic CYP3A4 abundance, albumin concentration, BMI, body surface area, sex; (f) hepatic CYP3A4 abundance, albumin concentration, BMI, body surface area, sex, weight; (g) hepatic CYP3A4 abundance, albumin concentration, BMI, body surface area, sex and weight<sup>[23]</sup>.

ผลจากการทำ stepwise forward inclusion โดยพิจารณาจากนัยสำคัญของตัวแปรที่ถูกคัดเลือกเข้าในแต่ละลำดับ ( $F \leq 0.05$ ) ต่อประสิทธิภาพการพยากรณ์ระดับของยา sorafenib ในกระแสเลือด พบว่าปัจจัยด้าน hepatic cytochrome P450 3A4 (CYP3A4) abundance ระดับ albumin และ body mass index (BMI) เป็นปัจจัยหลักที่มีอิทธิพลต่อความแปรปรวนของการตอบสนองของยา sorafenib (ตารางที่ 2) โดย

multivariable model ที่ประกอบด้วยทั้งสามตัวแปรดังกล่าวสามารถพยากรณ์ระดับของยา sorafenib ในกระแสเลือดได้ถูกต้องถึง 86.8% ในขณะที่เมื่อผนวกรวม body surface area, เพศ และน้ำหนักตัวเข้ากับ model ก่อนหน้า พบว่าประสิทธิภาพการพยากรณ์เท่ากับ 88.3% ซึ่งเพิ่มขึ้นเพียงเล็กน้อยเท่านั้น<sup>[23]</sup>

ผลของการศึกษานี้จะเป็นแนวทางเบื้องต้นในการพิจารณาการปรับขนาดของยา sorafenib ทางคลินิกโดยเป็นข้อมูลประกอบการตัดสินใจว่า หากผู้ป่วยมีโรคร่วมหรือปัจจัยที่ส่งผลให้เกิดความเปลี่ยนแปลงต่อตัวแปรที่มีผลต่อระดับของยา sorafenib ตามที่ได้ทำการ simulation การปรับขนาดของยาจึงควรกระทำโดยความระมัดระวังมากขึ้น หรือการตรวจค่าทางห้องปฏิบัติการเพิ่มเติมเพื่อประกอบการตัดสินใจการวางแผนการรักษาด้วยยา sorafenib ในลำดับถัดไป

### 3.3 บทสรุป

Precision medicine เป็นแนวปฏิบัติใหม่ที่มุ่งเน้นการบริหารยาให้เกิดความเหมาะสมที่สุดต่อผู้ป่วย โดยคำนึงถึงประเด็นความแตกต่างหลายมิติเข้าด้วยกัน ความก้าวหน้าด้านวิทยาการคอมพิวเตอร์และวิทยาการข้อมูลกำลังเข้ามามีบทบาทสำคัญอย่างยิ่งต่อการพัฒนาศาสตร์แขนงนี้และที่เกี่ยวข้อง การประยุกต์ใช้วิทยาการข้อมูลในทางเภสัชกรรม จึงเป็นแนวทางที่จะช่วยส่งเสริมให้เกิดการใช้ยามีประสิทธิภาพมากขึ้นและถือเป็นบทบาทที่ทำท้ายของเภสัชกรในการที่จะช่วยพัฒนาระบบสาธารณสุขของประเทศ

### 3.4 เอกสารอ้างอิง

1. Mass RD, Press MF, Anderson S, Cobleigh MA, Vogel CL, Dybdal N, Leiberman G, Slamon DJ (2005) Evaluation of clinical outcomes according to HER2 detection by fluorescence in situ hybridization in women with metastatic breast cancer treated with trastuzumab. *Clinical Breast Cancer* 6 (3):240-246. doi:10.3816/CBC.2005.n.026
2. Giordano SH, Temin S, Chandarlapaty S, Crews JR, Esteva FJ, Kirshner JJ, Krop IE, Levinson J, Lin NU, Modi S, Patt DA, Perlmutter J, Ramakrishna N, Winer EP, Davidson NE (2018) Systemic Therapy for Patients With Advanced Human Epidermal Growth Factor Receptor 2-Positive Breast Cancer: ASCO Clinical Practice Guideline Update. *Journal of Clinical Oncology* 36 (26):2736-2740. doi:10.1200/jco.2018.79.2697
3. Cystic Fibrosis Foundation. Cystic Fibrosis Foundation Patient Registry 2019 Annual Data Report., 2019 edn. Cystic Fibrosis Foundation, Bethesda



4. Wainwright CE, Elborn JS, Ramsey BW, Marigowda G, Huang X, Cipolli M, Colombo C, Davies JC, De Boeck K, Flume PA, Konstan MW, McColley SA, McCoy K, McKone EF, Munck A, Ratjen F, Rowe SM, Waltz D, Boyle MP (2015) Lumacaftor–Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del CFTR. *New England Journal of Medicine* 373 (3):220-231. doi:10.1056/NEJMoa1409547
5. Primorac D, Bach-Rojecky L, Vađunec D, Juginović A, Žunić K, Matišić V, Skelin A, Arsov B, Boban L, Erceg D, Ivkošić IE, Molnar V, Ćatić J, Mikula I, Boban L, Primorac L, Esquivel B, Donaldson M (2020) Pharmacogenomics at the center of precision medicine: challenges and perspective in an era of Big Data. *Pharmacogenomics* 21 (2):141-156. doi:10.2217/pgs-2019-0134
6. Chambliss AB, Chan DW (2016) Precision medicine: from pharmacogenomics to pharmacoproteomics. *Clinical Proteomics* 13 (1):25. doi:10.1186/s12014-016-9127-8
7. Weinshilboum RM, Wang L (2017) Pharmacogenomics: Precision Medicine and Drug Response. *Mayo Clinic Proceedings* 92 (11):1711-1722. doi:10.1016/j.mayocp.2017.09.001
8. Olivier C, Williams-Jones B (2011) Pharmacogenomic technologies: a necessary "luxury" for better global public health? *Globalization and Health* 7 (1):30. doi:10.1186/1744-8603-7-30
9. Séguin B, Hardy B-J, Singer PA, Daar AS (2008) Genomic medicine and developing countries: creating a room of their own. *Nature Reviews Genetics* 9 (6):487-493. doi:10.1038/nrg2379
10. El Shamieh S, Zgheib NK (2021) Pharmacogenetics in developing countries and low resource environments. *Human Genetics*. doi:10.1007/s00439-021-02260-9
11. Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JPA, Salman RA-S, Chan A-W, Glasziou P (2014) Biomedical research: increasing value, reducing waste. *The Lancet* 383 (9912):101-104. doi:10.1016/S0140-6736(13)62329-6
12. Ioannidis JPA (2014) How to Make More Published Research True. *PLoS Medicine* 11 (10):e1001747. doi:10.1371/journal.pmed.1001747
13. López-Rubio E, Ratti E (2019) Data science and molecular biology: prediction and mechanistic explanation. *Synthese*. doi:10.1007/s11229-019-02271-0

14. Navarro FCP, Mohsen H, Yan C, Li S, Gu M, Meyerson W, Gerstein M (2019) Genomics and data science: an application within an umbrella. *Genome Biology* 20 (1):109.  
doi:10.1186/s13059-019-1724-1
15. Boran ADW, Iyengar R (2010) Systems pharmacology. *The Mount Sinai journal of medicine, New York* 77 (4):333-344. doi:10.1002/msj.20191
16. Ladner JT, Grubaugh ND, Pybus OG, Andersen KG (2019) Precision epidemiology for infectious disease control. *Nature Medicine* 25 (2):206-211. doi:10.1038/s41591-019-0345-2
17. Fu Z, Zhang R, Li P, Jia M (2019) Translational epidemiology: The powerful tool for precision cancer medicine. *Journal of Cancer Research and Therapeutics* 15 (2):269-271.  
doi:10.4103/jcrt.JCRT\_276\_18
18. Saracci R (2018) Epidemiology in wonderland: Big Data and precision medicine. *European Journal of Epidemiology* 33 (3):245-257. doi:10.1007/s10654-018-0385-9
19. Prosperi M, Min JS, Bian J, Modave F (2018) Big data hurdles in precision medicine and precision public health. *BMC Medical Informatics and Decision Making* 18 (1):139.  
doi:10.1186/s12911-018-0719-2
20. Cheng AL, Kang YK, Lin DY, Park JW, Kudo M, Qin S, Chung HC, Song X, Xu J, Poggi G, Omata M, Pitman Lowenthal S, Lanzalone S, Yang L, Lechuga MJ, Raymond E (2013) Sunitinib versus sorafenib in advanced hepatocellular cancer: results of a randomized phase III trial. *Journal of Clinical Oncology* 31 (32):4067-4075. doi:10.1200/jco.2012.45.8372
21. Ruanglertboon W, Sorich MJ, Rowland A, Hopkins AM (2020) Effect of early adverse events resulting in sorafenib dose adjustments on survival outcomes of advanced hepatocellular carcinoma patients. *International Journal of Clinical Oncology* 25 (9):1672-1677. doi:10.1007/s10147-020-01698-7
22. Ruanglertboon W, Sorich MJ, Logan JM, Rowland A, Hopkins AM (2020) The effect of proton pump inhibitors on survival outcomes in advanced hepatocellular carcinoma treated with sorafenib. *Journal of Cancer Research and Clinical Oncology* 146 (10):2693-2697.  
doi:10.1007/s00432-020-03261-3

23. Ruanglertboon W, Sorich MJ, Hopkins AM, Rowland A (2021) Mechanistic Modelling Identifies and Addresses the Risks of Empiric Concentration-Guided Sorafenib Dosing. *Pharmaceuticals* 14 (5):389