CHAPTER 15

STRATEGY IN DRUG DISCOVERY AND DEVELOPMENT

"Science finds, industry applies, and man conforms."

—Anonymous

Biopharmaceutical modeling will be an effective tool to improve the productivity of drug discovery and development if we use it appropriately [1]. At present, several software packages are commercially available [2]. Biopharmaceutical modeling will also be useful for the quality-by-design strategy [3–6].

Drugs with low bioavailability tend to show variable Cp-time profiles [7]. (Fig. 15.1). Therefore, to increase the success rate of drug development, it is preferable to have a drug candidate with high bioavailability. Therefore, it would be preferable to have a strategy to design and select a candidate drug with appropriate BA% from the early stages of drug discovery.

15.1 LIBRARY DESIGN

The quality of a compound library directly affects the quality of a lead compound and may impact the quality of a clinical candidate compound and the success rate of drug development. Therefore, a lead compound with a reasonable biopharmaceutical profile had better be discovered from the compound library. Generally, during the lead optimization stage, the average solubility of a compound series decreases because the average of molecular weight (MW) and lipophilicity would increase to achieve a high pharmacological potency and selectivity [8, 9]. The

Biopharmaceutics Modeling and Simulations: Theory, Practice, Methods, and Applications, First Edition. Kiyohiko Sugano.

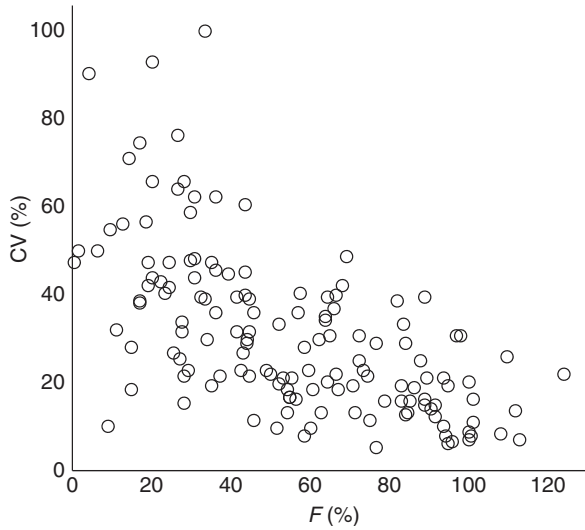

Figure 15.1 Relationship between absolute bioavailability (F) and intersubject variability (CV) in absolute bioavailability. *Source:* Replotted from Reference 7.

large number of hydrogen bonds and high lipophilicity might cause high crystalline energy and high hydrophobicity, leading to low solubility (Section 2.2.1).

In order to find a high quality lead compound, "drug-likeness" [10, 11] and structural diversity should be considered in library design. "Drug-likeness" can be assessed by a simple rule such as the "rule of five," which calculates the molecular weight, the number of hydrogen bond donors and acceptors, and the lipophilicity [10–12]. These factors also affect the biopharmaceutical profiles. High MW might cause either or both of ADME property and synthetic complexity issues and might lead to a development candidate with a low success rate. Therefore, it is preferred that the MW of the library compounds is set as low as possible (e.g., <400) [13]. In addition to oral absorption of a drug, lipophilicity also affects the volume of distribution [14–16], renal clearance (renal reabsorption) [17, 18], etc.

For library design, an *in silico* approach is ordinary (Section 5.1) [19]. However, at present, the prediction accuracy of *in silico* tools is not completely satisfactory. If one keeps in mind its limitations, it can be used for the purpose of library design.

Even though the concept of "drug-likeness" would be the baseline for library design, it should also be remembered that there are many exceptions (Fig. 15.2) [20]. It is important to be flexible so as not to fail the chance.

15.2 LEAD OPTIMIZATION

In the lead optimization stage, medium- to high throughput screening data will become available. An apparent solubility screening with PLM crystallinity

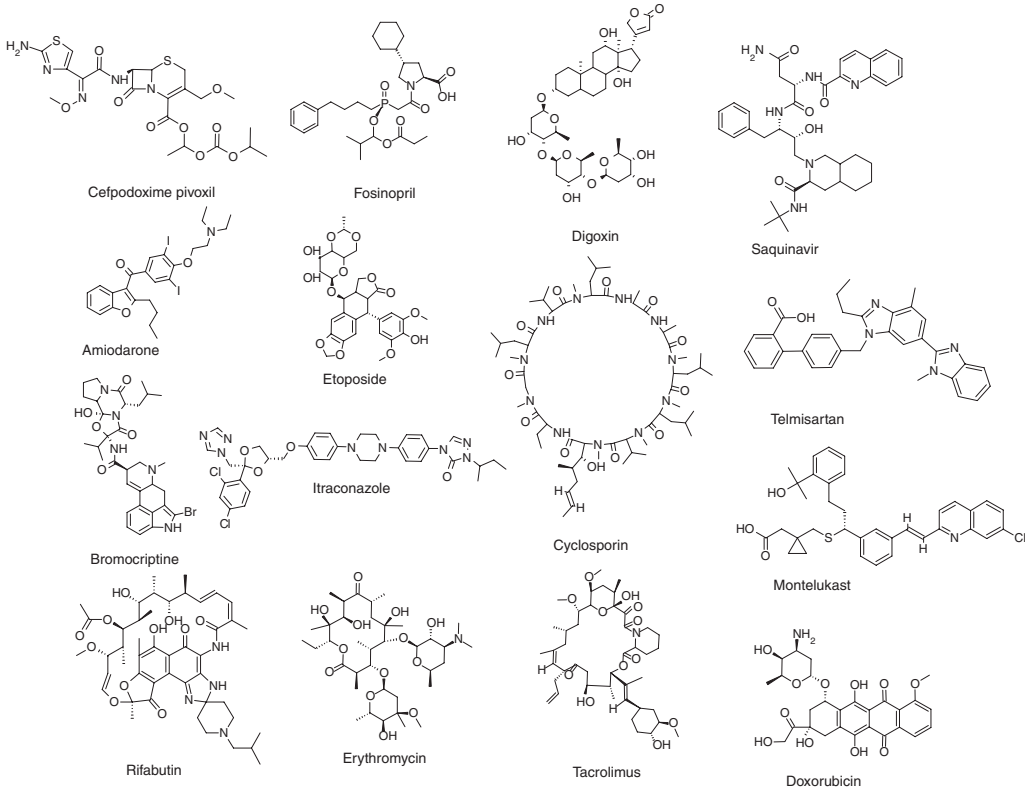

Figure 15.2 Marketed drugs outside the rule of five.

assessment [21] can be performed in parallel with a membrane permeability assay such as PAMPA. These information can be used for the initial assessment of the BCS category for the chemical series [22]. These data are usually stored into an in-house database. Biopharmaceutical modeling can be integrated to the in-house database and automatically run in the background.

At the late optimization stage, in addition to the basic physicochemical data, the solubility data in biorelevant media such as FaSSIF would be available.

15.3 COMPOUND SELECTION

Compounds that passed *in vitro* assays would then be evaluated by *in vivo* assays. Apparent solubility assay with powder material accompanied with crystallinity evaluation should take place at this stage. These data can be used to run detailed biopharmaceutical modeling. Biopharmaceutical modeling can be useful to select an appropriate formulation for *in vivo* studies (such as early toxicology studies) and also to interpret the *in vivo* results.

15.4 API FORM SELECTION

A detailed solubility profile (the pH–solubility profile and the effect of bile) should be studied at this stage as part of preformulation studies [23, 24]. The effect of bile should also be evaluated to assess the effect of food on oral absorption. If the candidate compounds (may be 1–3 compounds from a project) have low solubility, salt formation, cocrystal formation, and/or particle size reduction would be investigated to improve the dissolution profile. Miniscale dissolution tests can be used at this stage [25, 26].

In addition to the biopharmaceutical performance, developability of the API (stability, production suitability, production facility, ease of quality control, etc.) should be simultaneously evaluated [23, 24].

15.5 FORMULATION SELECTION

If the API form optimization was not successful to give sufficient exposure, particle size reduction would be the next measure. In the case of the dissolution-rate-limited absorption, particle size reduction may be effective to increase Fa%. The effectiveness of particle size reduction can be assessed by biopharmaceutical modeling. Usually, the initial dissolution rate is reciprocal to the particle size. It is worth mentioning that a milling process can change the solid form, especially to an amorphous state. Therefore, the milling feasibility should be simultaneously studied.

If the API form selection and standard particle size reduction (ca. 5–10 μm) was not successful to improve Fa%, special formulations such as nanoparticle, solid dispersion, and SEDDS could be the next option to achieve a target *in vivo* exposure. The mechanism-based flow chart (Fig. 11.18) may be a useful guide.

However, the successful development of special formulations is not always guaranteed. Therefore, we suppose that the primary solution to the low solubility issue would be to fix it in chemistry (compound structure or API form) [27]. If low solubility is inevitable, we should then challenge special formulations. To reduce the risk of the special formulations while exploring the possibilities, it is preferable to experimentally examine these special formulations as early as possible in the drug discovery and development process. At the same time as the assessment of oral absorption, developability and market competitiveness (compliance, cost of goods, development speed, etc.) should be addressed. The development of special formulations is always expensive with respect to time and man power. The maximum loading dose in special formulations is often smaller than that of the standard formulations. To avoid over and under expectation on special formulations, sufficient accountability for the project team is necessary. It is important to have a decision tree before the *in vivo* oral absorption study, because a discovery project team can be fascinated by the significant enhancement of oral absorption by a special formulation and developability and market competitiveness are put aside.

15.6 STRATEGY TO PREDICT HUMAN FA%

It is important to validate biopharmaceutical modeling by comparing with the real experimental data. A step-by-step cross validation strategy is shown in Figure 15.3. If there is a discrepancy between simulated and observed data,

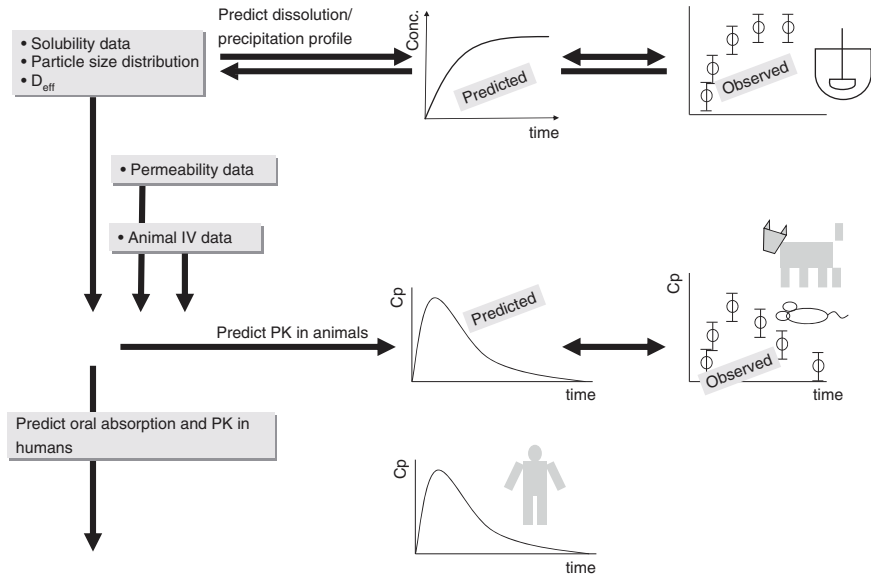

Figure 15.3 Step-by-step cross validation.

the reason for the discrepancy should be investigated by independent mechanistic experiments, rather than solely estimated by parameter fitting. "All-in-all" validation and compound by compound parameter optimization should not be taken [28]. The parameter responsible for the simulation error cannot be selected simply by try and error curve fitting. For example, the degree of flatness (DF) in the $P_{\text{eff}}-k_{\text{perm}}$ equation might be selected and optimized, even though the real reason was an error in P_{eff}. If this DF is used for humans and dose dependency is simulated, the prediction would fail.

REFERENCES

1. Kuentz, M., Nick, S., Parrott, N., Rothlisberger, D. (2006). A strategy for preclinical formulation development using GastroPlus(TM) as pharmacokinetic simulation tool and a statistical screening design applied to a dog study. *Eur. J. Pharm. Sci*., **27**, 91–99.
2. Parrott, N., Lave, T. (2002). Prediction of intestinal absorption: comparative assessment of GASTROPLUS and IDEA. *Eur. J. Pharm. Sci*., **17**, 51–61.
3. Jiang, W., Kim, S., Zhang, X., Lionberger, R. A., Davit, B. M., Conner, D. P., Yu, L. X. (2011). The role of predictive biopharmaceutical modeling and simulation in drug development and regulatory evaluation. *Int. J. Pharm*., **418**, 151–160.
4. Yu, L. X. (2008). Pharmaceutical quality by design: product and process development, understanding, and control. *Pharm. Res*., **25**, 781–791.
5. Lionberger, R. A., Lee, S. L., Lee, L., Raw, A., Yu, L. X. (2008). Quality by design: concepts for ANDAs. *AAPS J*., **10**, 268–276.
6. Henck, J. O., Byrn, S. R. (2007). Designing a molecular delivery system within a preclinical timeframe. *Drug Discovery Today*, **12**, 189–199.
7. Hellriegel, E. T., Bjornsson, T. D., Hauck, W. W. (1996). Interpatient variability in bioavailability is related to the extent of absorption: implications for bioavailability and bioequivalence studies. *Clin. Pharmacol. Ther*., **60**, 601–607.
8. Sugano, K., Obata, K., Saitoh, R., Higashida, A., Hamada, H., Processing of Biopharmaceutical Profiling Data in Drug Discovery, in: B. Testa, S., Krämer, H., Wunderli-Allenspach, G., Folkers (eds) *Pharmacokinetic Profiling in Drug Research*, Wiley-VCH, Zurich, 2006, pp. 441–458.
9. Wenlock, M. C., Austin, R. P., Barton, P., Davis, A. M., Leeson, P. D. (2003). A comparison of physiochemical property profiles of development and marketed oral drugs. *J. Med. Chem*., **46**, 1250–1256.
10. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev*., **23**, 3–25.
11. Lipinski, C. A. (2004). Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today Technol*., **1**, 337–341.
12. Leeson, P. D., Springthorpe, B. (2007). The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discovery*, **6**, 881–890.
13. Oprea, T. I., Davis, A. M., Teague, S. J., Leeson, P. D. (2001). Is there a difference between leads and drugs? A historical perspective. *J. Chem. Inf. Comput. Sci*., **41**, 1308–1315.

14. Lombardo, F., Obach, R. S., Shalaeva, M. Y., Gao, F. (2002). Prediction of volume of distribution values in humans for neutral and basic drugs using physicochemical measurements and plasma protein binding data. *J. Med. Chem*., **45**, 2867–2876.
15. Rodgers, T., Rowland, M. (2006). Physiologically based pharmacokinetic modelling 2: predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. *J. Pharm. Sci*., **95**, 1238–1257.
16. Rodgers, T., Rowland, M. (2007). Mechanistic approaches to volume of distribution predictions: understanding the processes. *Pharm. Res*., **24**, 918–933.
17. Barza, M., Brown, R. B., Shanks, C., Gamble, C., Weinstein, L. (1975). Relation between lipophilicity and pharmacological behavior of minocycline, doxycycline, tetracycline, and oxytetracycline in dogs. *Antimicrob. Agents Chemother*., **8**, 713–720.
18. Varma, M. V., Feng, B., Obach, R. S., Troutman, M. D., Chupka, J., Miller, H. R., El-Kattan, A. (2009). Physicochemical determinants of human renal clearance. *J. Med. Chem*., **52**, 4844–4852.
19. Lombardo, F., Gifford, E., Shalaeva, M. Y. (2003). In silico ADME prediction: data, models, facts and myths. *Mini-Rev. Med. Chem*., **3**, 861–875.
20. Sakaeda, T., Okamura, N., Nagata, S., Yagami, T., Horinouchi, M., Okumura, K., Yamashita, F., Hashida, M. (2001). Molecular and pharmacokinetic properties of 222 commercially available oral drugs in humans. *Biol. Pharm. Bull*., **24**, 935–940.
21. Sugano, K., Kato, T., Suzuki, K., Keiko, K., Sujaku, T., Mano, T. (2006). High throughput solubility measurement with automated polarized light microscopy analysis. *J. Pharm. Sci*., **95**, 2115–2122.
22. Obata, K., Sugano, K., Machida, M., Aso, Y. (2004). Biopharmaceutics classification by high throughput solubility assay and PAMPA. *Drug Dev. Ind. Pharm*., **30**, 181.
23. Fiese, E. F. (2003). General pharmaceutics-the new physical pharmacy. *J. Pharm. Sci*., **92**, 1331–1342.
24. Balbach, S., Korn, C. (2004). Pharmaceutical evaluation of early development candidates "the 100mg-approach". *Int. J. Pharmaceutics*, **275**, 1–12.
25. Takano, R., Sugano, K., Higashida, A., Hayashi, Y., Machida, M., Aso, Y., Yamashita, S. (2006). Oral absorption of poorly water-soluble drugs: computer simulation of fraction absorbed in humans from a miniscale dissolution test. *Pharm. Res*., **23**, 1144–1156.
26. Avdeef, A., Tsinman, O. (2008). Miniaturized rotating disk intrinsic dissolution rate measurement: effects of buffer capacity in comparisons to traditional wood's apparatus. *Pharm. Res*., **25**, 2613–2627.
27. Lipinski, C. (2004). Solubility in the design of combinatorial libraries. *Chem. Anal*., **163**, 407–434.
28. Jones, H. M., Parrott, N., Jorga, K., Lave, T. (2006). A novel strategy for physiologically based predictions of human pharmacokinetics. *Clin. Pharmacokinet*, **45**, 511–542.

CHAPTER 16

EPISTEMOLOGY OF BIOPHARMACEUTICAL MODELING AND GOOD SIMULATION PRACTICE

> "The greatest obstacle to discovering the shape of the earth, the continents and the ocean was not ignorance but the illusion of knowledge."
>
> —Daniel J. Boorstin

16.1 CAN SIMULATION BE SO PERFECT?

In the literature, accurate predictions of the C_p–time profile by using a commercial software package have been often advocated. However, as discussed in Section 8.1, considering the uncertainty and variations in the input data and the model equations (as well as the variation in *in vivo* data), such accurate predictions should be unattainable.

It is often assumed without judgment that a commercial software package should have been fully validated. However, it would be a good practice to investigate the software package before using it. Appropriateness of the equations should be thoroughly investigated, for example, equations for the solid surface pH, the common ion effect, the paracellular pathway, the pH-partition theory with microclimate pHs, the UWL permeation, the nucleation, and the unbound fraction. In addition, appropriateness of physiological parameters should be thoroughly investigated, for example, the intestinal fluid volume and the degree of

Biopharmaceutics Modeling and Simulations: Theory, Practice, Methods, and Applications, First Edition. Kiyohiko Sugano.

flatness (or SA_{GI}/V_{GI} ratio). In several reports, the intestinal fluid volume of 600–1500 ml was used (should be 100–250 ml) and the SA_{GI}/V_{GI} ratio was set to be ca. 1.3 (should be ca. 2.3).[1] Furthermore, the drug data used in a report should also be carefully checked. All experimental scientists know that the results of an *in vitro* assay have variations, which often become more than twofold. Artificial intestinal fluids (such as FaSSIF) and *in vitro* membrane assays (such as Caco-2) cannot be a perfect surrogate for corresponding *in vivo* factors. Incorrect drug parameters were sometimes used in the literature. For example, in one report, instead of the true density of the drug, the tap density was used for surface area calculation. In another report, a pK_a value for an acid was used for a base. In many cases, because the details of simulation were not fully disclosed, it is not clear why accurate simulation results were obtained despite the insufficiency of equations and/or the use of incorrect input data.

As discussed in Chapter 8, even for the simplest cases, the current average simulation error of the GUT framework (all of which is constructed by public scientific knowledge), is ca. twofold. A scientific progress of this area should start with admitting this reality.

16.2 PARAMETER FITTING

Probably, drug-by-drug parameter fitting is one of the most frequent reasons to give a superficially good simulation.

Parameter fitting is sometimes performed unwittingly. For example, when we have four different P_{eff} values estimated from the PAMPA, Caco-2, MDCK, rat *in situ* perfusion, and *in silico* methods, we might select one method, on a drug-by-drug basis, which gives the best fitting to *in vivo* results. This is mathematically equivalent to doing a parameter fitting for each drug. If we were to select a suitable assay, it should be based on an independent reason (e.g., because the drug is a transporter substrate, rat *in situ* perfusion data is used). In some cases, as clinical i.v. data was not available, CL and Vd were obtained by fitting to p.o. data. In one report, different CL and Vd values were used for i.v. and p.o. simulations, respectively.

However, parameter fitting is sometimes necessary when a physiological parameter cannot be directly obtainable and have to be back-estimated from the clinical PK data of drugs. For this purpose, the drugs that are free from

[1] In some reports, $V_{GI} = 600$ ml was used with the surface area of 800 cm^2, that is, $SA_{GI}/V_{GI} = 1.3$, which is equal to the cylindrical tube shape. Compared to the current most credible values of $V_{GI} = 130$ ml and $SA_{GI}/V_{GI} = 2.3$, the previous V_{GI} is larger and the previous SA_{GI}/V_{GI} is smaller. These errors worked in opposite directions and were coincidently canceled out, resulting in semiquantitative Fa% prediction for solubility-permeability-limited cases. However, with $V_{GI} = 600$ ml, the inflation point in the dose–AUC curve would be upshifted (Fig. 10.1). In addition, with $SA_{GI}/V_{GI} = 1.3$, for permeability-limited cases, human Fa% is underestimated by ca. twofold from the experimental P_{eff} values in humans (Fig. 8.2).

the uncertainty in the other factors should be used for parameter fitting. The number of model drugs should be sufficient to avoid overlearning (>5–10 data points per parameter). From a single C_p–time profile, only a little information is identifiable [1]. For example, the one-compartment model with three parameters, that is, k_a, k_{el}, and Vd, is usually sufficient to describe the oral PK profile of a drug. On the other hand, a mechanistic biopharmaceutical model contains dozens to hundreds of parameters. It is often difficult to identify a correct parameter for optimization solely from the C_p–time profile (Section 15.6).

16.3 GOOD SIMULATION PRACTICE

In modeling and simulation, transparency is a paramount requisite [2]. We have to exert every effort to improve the transparency of simulation processes. It is the cost we must pay for a healthy development of sciences.

16.3.1 Completeness

All the model equations and physiological parameters used in a simulation should be fully disclosed in a report or appropriate references should be cited so that independent readers can check the report. It is often the case that only the name of a commercial software package is described and mentioned as "the default setting was used." In this case, it is often difficult for the readers to judge the scientific rigor,[2] especially for the ones who do not have access to the commercial software package. The details of mechanistic equations and physiological parameters are usually described in the user's manual. However, the user's manual is not disclosed for public readers (even for journal referees). The transparency should be provided not only for the users but also for the public readers.

In addition, experimental conditions to obtain the drug parameters should be fully described in a report or appropriate references should be added Table 16.1. The API information such as solid form (free/salt, crystalline/amorphous, hydrate/anhydrate) and particle size should be reported. When simulating the C_p–time profile, the method to obtain CL, Vd, Fg, and Fh should be fully described.

A failed simulation should be reported. It is not a failure but a clue for progresses in the future. When parameter optimization is performed,[3] the simulation results before and after optimization should be reported.

[2]This is the reason why publications using commercial software packages are not used as scientific references in this book.

[3]As discussed above, drug-by-drug parameter optimization is not recommended.

16.3.2 Comprehensiveness

When a complicated model is used, the essence of the biopharmaceutical profile of a drug can be lost in complicated description. Even if all the simulation details were disclosed, it would be practically impossible for a third party to trace all simulation processes. To increase the comprehensiveness, the dose number, the dissolution number, and the permeation number should be at least reported. These dimensionless parameters can be used to capture the regime of oral absorption of a drug. This helps us to focus on the most important part of biopharmaceutical modeling of the drug. The use of a simpler model should be considered when it is sufficient (the Occam's razor, or parsimony principle). Even when all the factors are automatically calculated by a program, it would be helpful to describe which factor is/is not important. For example, a description like "because MW = 600, the contribution of the paracellular pathway is negligible" would be helpful for readers. When showing C_p–time profiles for the purpose of investigating the absorption phase, a log-normal plot should not be used. The use of support lines in a figure (Fig. 16.1) should be kept minimal. Even when calculating the C_p–time profile, Fa% data should be also reported.

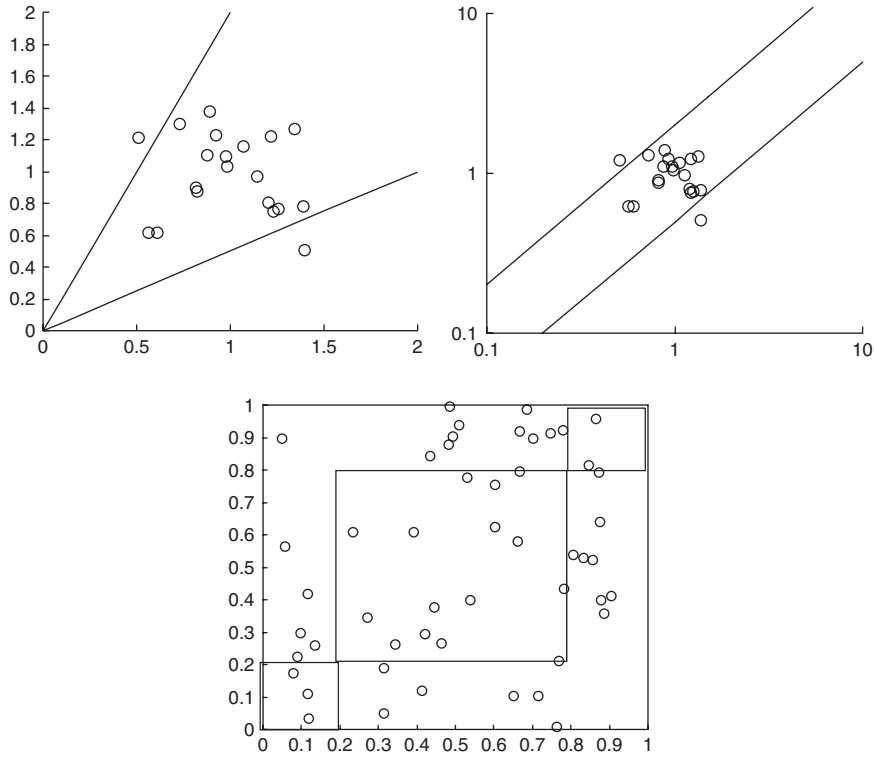

Figure 16.1 Types of support lines.

TABLE 16.1 List of Drug Parameters to be Reported (An Example)

Data	Comments
MW	Free form/salt/cocrystal/hydrate
Chemical structure	If possible to disclose
Chemical formula	If possible to disclose
pK_a	Experimental value is preferable
log P_{oct}	Experimental value is preferable
API solid form	Free form/salt/cocrystal/hydrate Crystalline/amorphous [lot number (for internal report)]
Particles	Size (D50, D90, SD, etc.) Shape
True density	Not tap and bulk density
Solubility	Final pH
	Final solid form
	Buffer species
	Bile micelle composition and concentration
	Solid separation method (filtration, centrifuge)
	Quantification method (HPLC, LC-MS, UV, etc.)
Permeability	Membrane type (Caco-2, MDCK, PAMPA, etc)
	pH
	Other additives in the media (bile micelles, BSA, etc)
	Agitation condition (or UWL thickness) for $P_{app} > 20 \times 10^{-6}$ cm/s cases
	Validity indicator (TEER, permeability standards (e.g., metoprolol))

REFERENCES

1. Yates, J.W., Jones, R.D., Walker, M., Cheung, S.Y. (2009). Structural identifiability and indistinguishability of compartmental models. *Expert Opin. Drug Metab. Toxicol*., **5**, 295–302.
2. Williams, P., Kim, Y., Ette, E., The epistemology of pharmacometrics, in: *Pharmacometrics*, P. Williams, E. Ette (Eds.) John Wiley & Sons, Inc., Hoboken, New Jersey, 2007, pp. 223–244.