Citation: CPT Pharmacometrics Syst. Pharmacol. (2017) 6, 727–729; doi:10.1002/psp4.12232

PERSPECTIVE

Improving Realism in Clinical Trial Simulations via Real-World Data

Holly Kimko[1]* and Kwan Lee[1]

Simulation validity depends on how well sampling distributions used reflect real-patient characteristics, such as drug adherence, disease progression, and pharmacologic handling in the body. We challenge the current use of growth charts from nondisease-specific pediatrics in simulations for drug development. Complementary use of data from clinical trials and the real-world is expected to achieve a more realistic representation of clinical outcomes for decisions in drug development, regulatory approval, and health technology assessment.
CPT Pharmacometrics Syst. Pharmacol. (2017) **6**, 727–729; doi:10.1002/psp4.12232; published online 19 September 2017.

Clinical trial simulation has been a frequently used tool for drug development in various disease areas.[1] Clinical trial simulation generates clinical responses of virtual subjects by approximating (a) human behavior, (b) disease progress, and (c) drug behavior within the framework of a proposed clinical trial design using mathematical models and numerical methods. Human behavior includes trial execution characteristics, such as adherence in drug administration and missing records. Disease status may change during a trial, for which a disease progress model needs to be developed. Drug behavior in the body is generally characterized by pharmacokinetic and pharmacodynamic models. The clinical trial design in each simulation scenario includes dosage regimens, subject enrollment criteria, number of arms, number of subjects, and so forth.

For clinical trial simulations, researchers use random sampling techniques from a database to generate clinical outcomes for a large number of virtual subjects, which facilitates strategic decisions, such as study operating characteristics or go/no-go decisions. For simulations to be realistic, the source databases need to exhibit sampling distributions that reflect real-world outcomes. To illustrate this, we consider a current practice in pediatric clinical trial simulation.

In the field of pharmacometrics, researchers frequently use[2] the growth chart database from the Centers for Disease Control and Prevention (CDC; https://www.cdc.gov/growthcharts/). The CDC recommends healthcare providers to:

- Use the World Health Organization growth standards to monitor growth for infants and children ages 0–2 years old in the United States.
- Use the CDC growth charts for children age 2 years and older in the United States.

Both growth charts include statistics of growth-related variables, such as weight and height, which were compiled from national health examination surveys and supplemental data. It should be noted that the simulated data from the World Health Organization and CDC growth charts do not necessarily represent the growth of pediatrics with a certain disease of interest, but rather all the pediatrics regardless of their health status. In actual pediatric clinical trials, we usually enroll patients with a disease of interest as subjects. Therefore, if a disease condition alters the overall weight distribution of such patients in an age range, the clinical outcomes from actual trials may be different from those predicted using the growth charts.

What is real-world data?

The consortium under the Innovative Medicines Initiative defines real-world data (RWD) as "an overarching term for data on the effects of health interventions (such as benefits, risks, or resource use) that are not collected in the context of conventional randomized controlled trials. Instead, RWD are collected both prospectively and retrospectively from observations of routine clinical practice."[3] Real-world evidence (RWE) refers to the information on health care that is derived from multiple RWD sources outside typical clinical research settings. The major difference between RWE and other types of evidence is the data collection settings, such as clinical care, and home or community settings, as opposed to research-intensive settings, like randomized clinical trials.

RWE is derived from curating, standardizing, and analyzing RWD to obtain reliable information, which could inform all phases of drug discovery and development. However, it has mainly been used to answer questions in early drug development and postmarketing studies for safety surveillance or effectiveness comparison, whereas clinical development and regulatory review have used data collected from well controlled efficacy trials, which may differ from the real-world due to careful attention toward demonstrating efficacy and safety of an interventional drug. Therefore, it is important to recognize that RWD has the potential to enhance the knowledge gained from conventional clinical trials, although it may need adjustment for a "shelf-life" as the standard of care treatment used during the collection of RWD changes over time or for tentative changes in patients' physiological and/or clinical status with treatments.[4,5]

A cursory example of applying disease-specific data to clinical trial simulations

Due to the difficulty in enrolling pediatric subjects in clinical trials, regulatory agencies tend to be more receptive toward

[1]Quantitative Sciences, Janssen R&D, LLC, Spring House, Pennsylvania, USA. *Correspondence: H Kimko (HKimko@its.jnj.com)
Received 31 July 2017; accepted 9 August 2017; published online on 19 September 2017. doi:10.1002/psp4.12232

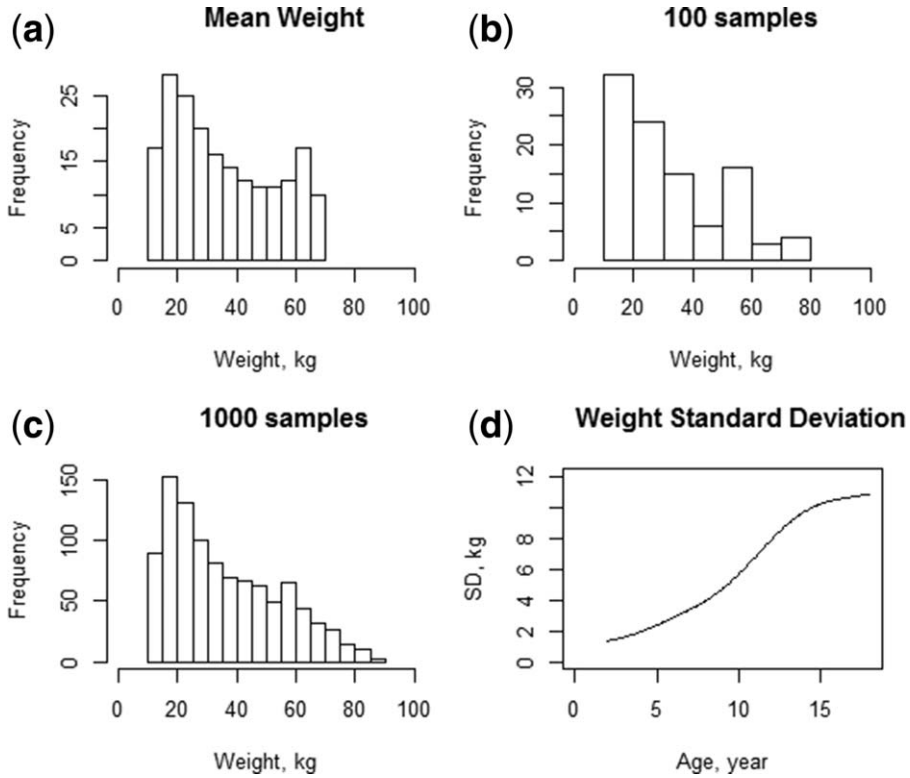

Figure 1 Weight distributions from the Centers for Disease Control and Prevention (CDC) weight growth chart for boys from 2–18 years old footnote: (**a**) distribution of recorded mean weights from the CDC growth chart database; (**b, c**) distribution for 100 and 1,000 randomly sampled subject's weights, respectively; and (**d**) SD of weight based on age.

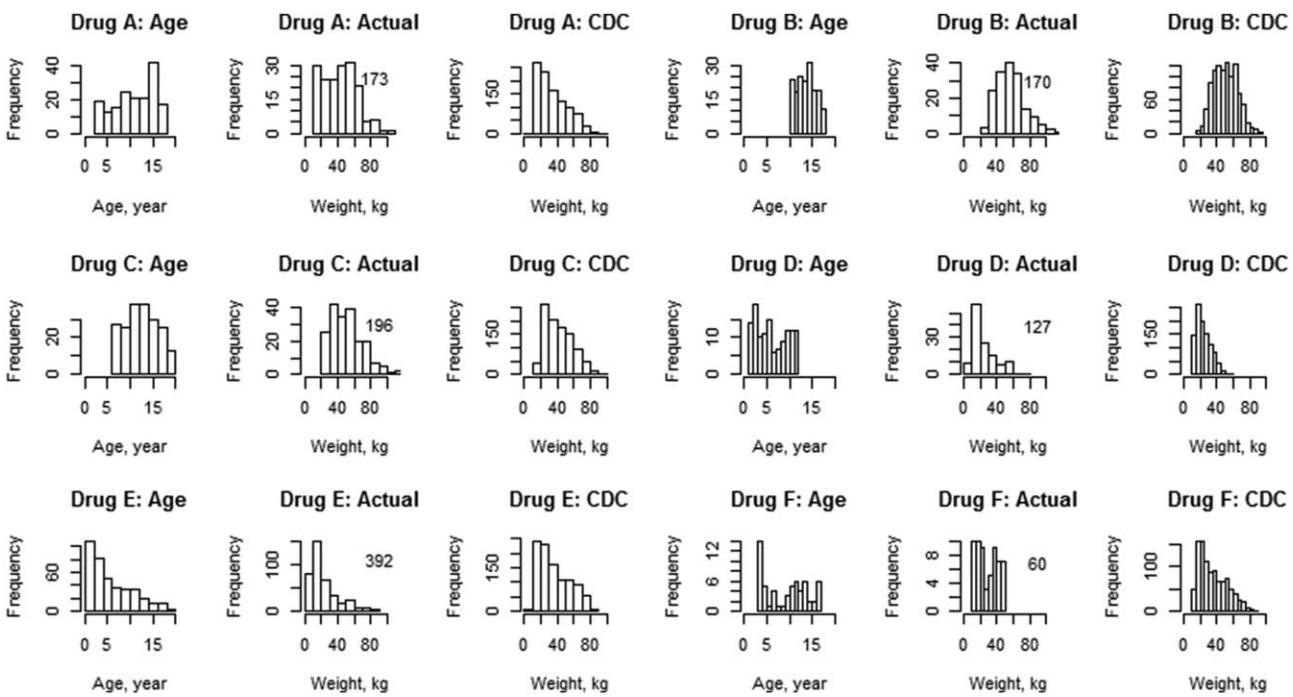

Figure 2 Age and weight distributions of pediatric studies with six different disease indications footnotes: drug A-F: juvenile rheumatoid arthritis, schizophrenia, epilepsy, gastric acid reflux, infection, and AIDS, respectively; "Actual" and Centers for Disease Control and Prevention ("CDC") indicate the distributions from the actual studies and the CDC growth chart, respectively. The numbers in the "Actual" plots are the numbers of subjects.

considering simulation results as a part of the totality of evidence for regulatory decisions. During simulations of pharmacokinetics of pediatrics, in which the pharmacokinetic model often includes body weight as a covariate, a certain number of ages are usually sampled uniformly, assuming the probability of enrolling pediatrics in each age group is the same, although it may not be realistic due to the correlation between disease type and age distribution. Then, a weight is sampled from the growth chart distribution at each sampled age. The weight-for-age percentile curves in the CDC growth chart (http://www.cdc.gov/growthcharts/data/set1clinical/cj41cs021bw.pdf) show that the weight distribution in younger pediatrics has a lower spread (SD; **Figure 1**) than that of older pediatrics. Therefore, lower weights in younger pediatrics have a greater chance to be resampled than higher weights in older pediatrics, yielding a right-skewed weight distribution after sampling for all pediatrics in the simulated clinical study, as shown in the weight distributions from 100 and 1,000 simulated subjects from the CDC weight growth chart for boys from 2–18 years old, using R software version 3.3.3 (**Figure 1**). The similar trend also exists for girls from 2–18 years old (data not shown).

In reality, the weight distribution of the subjects enrolled in actual clinical trials is dependent on the type of the diseases of interest. **Figure 2** shows the differences in distributions of age and weight from various pediatric studies in disease areas, such as juvenile rheumatoid arthritis, schizophrenia, epilepsy, gastric acid reflux, infection, and AIDS (drug A–F, respectively). A thought experiment to compare pharmacokinetic differences by sampling from a disease-nonspecific covariate distribution (e.g., the CDC growth chart) and a disease-specific distribution (whose shape is different) leads to the realization of an unrealistic representation of clinical study outcomes that use the CDC growth chart. Pediatric patients with AIDS often have impaired growth and maturation, resulting in lower weight and height than would be common for their age groups, which is shown in the case of drug F. Such patients may also have atypical laboratory values that impact the pharmacokinetics of various drugs as well, which calls for a careful selection of demographics and laboratory information to use in simulations.

The age and weight distribution examples shown in **Figure 2** are all from internal company databases, which were from studies with restricted subject inclusion and exclusion criteria. This type of "controlled" data from disease-specific subjects still does not accurately reflect the real-world, in which many additional factors, such as comedication, adherence, and persistence problems may yield effectiveness that is different from the efficacy demonstrated in well controlled trials. There are a few repositories of pediatric patient-specific data in real-world settings,[3] such as the Pediatric Health Information System through the Child Health Corporation of America (https://www.childrenshospitals.org/programs-and-services/data-analytics-and-research/pediatric-analytic-solutions/pediatric-health-information-system) and the Lexicomp Online internet-based drug information platform for multi-user groups in a networked system (http://webstore.lexi.com/ONLINE).

CONCLUSIONS

The RWD have been used in modeling and simulation of disease progression, patient care, and comparative effectiveness research.[6–8] Utilizing RWD to predict real-world effectiveness with efficacy data from randomized, controlled trials for drug approval dossiers or health economic evaluations seems to still be in its infancy stage, which is supported by the limited number of publications on this subject in medical literature.[9] Internal company data collected previously in a specific disease franchise may be used to support the efficacy of an investigational drug, but due to restrictions in enrollment inclusion/exclusion criteria and the investigator's intention to collect data with high integrity for regulatory submission purpose, such collected data often fail to represent effectiveness of the investigational drug. To improve realism in clinical trial simulations, it is recommended to use RWD that is accessible from enterprise-licensed databases (e.g., Optum, Truven); patient and population surveys; patient chart reviews; observational data from cohort studies; pragmatic clinical trials; and registries, etc.[10] By integrating RWD into clinical trial simulations, simulation results become more realistic in the sense that they are more generalizable in clinical practice settings. We expect more complementary use of data from controlled trials and the real-world in the near future.

Conflict of Interest. The authors declared no conflicts of interest.

1. Kimko, H.H.C. & Peck, C.C., eds. In *Clinical Trial Simulations: Applications and Trends*. AAPS Advances in the Pharmaceutical Sciences Series 1 (Springer, New York, NY, 2011).
2. Barrett, J.S. Modeling and Simulation in Pediatric Research and Development. In *Clinical Trial Simulations: Applications & Trends* (eds. Kimko, H.H.C. & Peck, C.C.) 401–433. AAPS Advances in the Pharmaceutical Sciences Series 1 (Springer, New York, NY, 2011).
3. RWE Navigator. Putting real-world healthcare data to work. <https://rwe-navigator.eu/homepage/about/>. Accessed 16 June 2017.
4. Booth, C.M. & Tannock, I.F. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. *Br. J. Cancer* **110**, 551–555 (2014).
5. Sherman, R.E. *et al.* Real-world evidence — What is it and what can it tell us? *N. Engl. J. Med.* **375**, 2293–2297 (2016).
6. Pobiruchin, M., Bochum, S., Martens, U.M., Kieser, M., Schramm, W. A method for using real world fata in breast cancer modeling. *J. Biomed. Inform.* **60**, 385–394 (2016).
7. Hughes, D. *et al.* Methods for integrating medication compliance and persistence in pharmacoeconomic evaluations. *Value Health* **10**, 498–509 (2007).
8. Wilke, T. *et al.* How to use pharmacy claims data to measure patient nonadherence? The example of oral diabetics in therapy of type 2 diabetes mellitus. *Eur. J. Health. Econ.* **14**, 551–568 (2013).
9. Panayidou, K. *et al.* GetReal in mathematical modelling: a review of studies predicting drug effectiveness in the real world. *Res. Synth. Methods.* **7**, 264–277 (2016).
10. Annemans, L., Aristides, M. & Kubin, M. Real life data: a growing need. ISPOR Connections. <http://www.ispor.org/news/articles/oct07/rld.asp>. Accessed 16 June 2017.