Evidence from Direct-To-Subject Study Designs for Health (Outcomes) Research and (Pharmaco) Epidemiology

Florian Eichmann, PhD, Principal Late Stage, Scientific Affairs and Real World Evidence, inVentiv Health Clinical, Munich, Germany

KEY POINTS . . .

Direct-to-subject research collects information directly from the research subject and provides valuable disease and treatment pattern information for Public Health, communication, or reimbursement purposes.

Active Surveillance and Pregnancy Registries are exemplary direct-to-subject study designs meeting regulatory post marketing requirements.

Validating outcomes is a primary quality aspect for subject based research, as demonstrated in a Design/Quality Matrix (DQM).

Introduction

Regulators and Health Technology Assessment (HTA) agencies - and thus manufacturers and research stakeholders - are showing increasing interest in real world information, optimally from the patient perspective, collected with high evidence validity, and readily available for analysis. This has led to competing health research trends (Fig. 1).

'Big Data' are often based on routine information collected by physicians, such as Electronic Medical Records (EMR) or administrative claims. In contrast, *'Patient Centered'* information comes directly from the research subject and includes subjective information that is not available from EMR or claims. High *'Evidence Value'* typically requires randomized, controlled, and thus experimental study setups. In contrast, *'Real World'* evidence should derive from routine, real life, non-experimental designs ensuring high external validity.

To meet stakeholders' requirements, and to save time and money, we must identify the optimal study design for a given research question. Direct-to-subject studies collect clinical, subject-reported, or economic information *directly* from the research subject: Instead of relying on indirect data sources such as health care providers or administrative databases, direct-to-subject studies, virtual trials, or subject-based research provide direct access to individual health information from the subject's perspective.

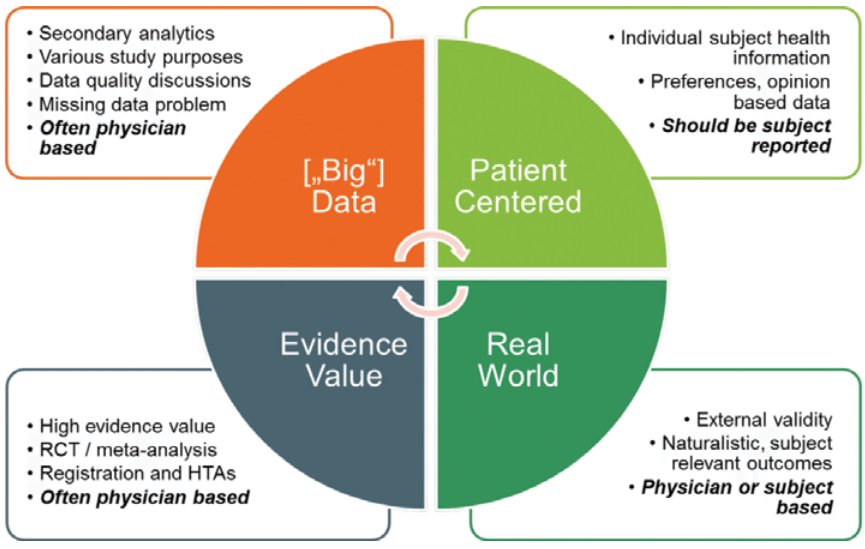
Regulatory and reimbursement stakeholders increasingly require subjects' perspectives to be included into the evidence portfolio supporting their decisions. Nevertheless, direct-to-subject studies are still falling behind relative to database approaches or site based trials. This article outlines strengths of subject-based research and indicates how to enhance direct-to-subject designs to maximize real world evidence validity.

Design/Quality Matrix

To assess study designs in a structured way, we are using a Design/Quality Matrix (DQM). The matrix lists study designs horizontally and quality factors vertically and depicts how study designs perform with regard to individual quality factors. We are focusing our evaluation on following study design 'archetypes':

- site-based interventional randomized trial (RCT)
- site-based observational primary research (Non-Interventional Study, Registry)

Figure 1: Competing Health Research Trends.



■ site-based observational database research (EMR, claims)
■ subject-based observational primary research (survey)
■ subject-based database research (e.g. 'National Health and Wellness Survey', 'Patients Like Me').

Research quality is a multifactorial, complex construct depending on operational and administrative factors and influenced by study objectives. Which quality factors should thus be considered? Validity of the study sample and thorough information on confounders and outcomes are of primary importance for a cross-sectional, one-time data collection. In longitudinal studies, loss to follow-up ('attrition') and availability of confounders and outcomes over the whole data collection period provide additional challenges. In summary, following main factors determine evidence value and quality of a health research study and thus form the vertical axis of our Design/Quality Matrix (DQM):

■ representative sample included (for disease, for product use)
■ relevant "independent variables / confounders" measurable
■ administrative quality, ethical, regulatory acceptability of study design and setup
■ attrition (loss to follow-up) low (for prospective longitudinal study)
■ outcomes of interest available and measurable (Clinical/Economic/Patient Reported Outcomes/Preferences…)

We will now look at exemplary research questions from both regulatory and non-regulatory areas and assess case study designs using the DQM.

Subject-based Public Health case study including outcome validation

To assess the influence of life style related risk factors during pregnancy on premature childbirth, a representative sample of pregnant women was recruited. Importantly, the study collected life style information prior to outcome information and thus independently from the study outcome.

Subjects were engaged in the study, they took care of aligning all data sources. Thus, attrition of engaged and participating subjects was low (87% retained to outcome 1 / 82% retained to outcome 2). Combining different data sources – including valid, i.e. physician confirmed, outcomes - improved outcome validity and thus evidence quality.

The most important lifestyle factors influencing premature birth were found to be: unemployed partner, mother living single and smoking. As a side effect of the study, radiation exposure had no measureable effect on premature childbirth [1,2].

Risk factors for premature birth: Linking subjective and medical data

Data content and variables	**Risk factors during pregnancy:** e.g. social stress, nutrition, smoking, alcohol, excercise, BMI, medical history, unemployment	**Prenatal medical screening & examinations:** screenings, e.g. diabetes, high blood pressure, x-ray, ultrasound, woman's health, medication, hereditary risk factors	**Perinatal birth data:** e.g. date of birth and length of pregnancy, type of birth, weight, height, asphyxia; Apgar score
Source of data	Self-completed questionnaire (woman)	Official maternity log: „Mutterpass" (gynecologist)	Medical record/ birth documentation (obstetrician)
Time period	Pregnancy weeks 15 - 28	During pregnancy; fixed intervals	At birth
N	3.946 →	3.429 →	3.218

Public Health Research

The most straightforward access to subject-based research is through public health or health services research. Public Health Research stakeholders routinely use social and market research instruments to deliver valid subject-centered study approaches. For many health research questions, health care provider sites do not yield appropriate data, and databases aggregating data collected by physicians often fail to contain relevant information.

Let's first look at a public health research case study scenario to assess advantages and disadvantages of study designs. The research question is: *How to investigate the influence of life style related risk factors during pregnancy on premature childbirth?*

Looking at the DQM for this case study (Table 1), an interventional trial will not be feasible as it would require artificially exposing pregnant women to life style risk factors. Site- and subject-based databases do not cover exposure to lifestyle factors and thus are not useful in this specific case.

As neither the interventional nor the database approaches will work for our research question, the observational primary research designs remain for consideration. Compared to a prospective or retrospective site based data collection, a subject based primary research design provides advantages with regard to access >

Table 1: Design/Quality Matrix (DQM): Public Health Research Case Study

Research Question: *Assess influence of life style related risk factors during pregnancy on premature childbirth*					
	Site Trial	Site Primary	Site Database	**Subject Primary**	**Subject Database**
	RCT…	OBS/NIS	EMR / Claims	PH / Survey	NHWS / plm ..
Sample / Epi Validity	?	(?)	(++)	++	
Independent variables / confounders		?	??	++	?
Admin Quality / Ethics / Reg /	--				
Attrition				++	
Outcome of interest		++		??→++	(?)

to independent variables (life style risk factors) and sample validity. The crucial question with many subject-based designs, however, is: How to measure outcomes, in this case pregnancy related outcomes, with appropriate evidence quality (see bottom row ?? in the DQM, Table 1).

Combining a subject-based study design with 'official' maternity logs and 'valid' physician based chart information (see Public Health case study box for case study design and results) 'validates' outcomes and thus generates a high quality study design (see bottom row ?? in Table 1 turn into a ++).

Regulatory Post Marketing Studies

Subject-centered approaches help fulfilling post marketing requirements imposed by regulators such as the FDA and EMA, and thus meet regulatory purposes. Let's now look at a regulatory example: *Authorities ask whether a newly-approved intrauterine contraceptive device (IUD) is associated with a higher risk of pelvic inflammatory disease compared to establish IUDs.*

In this scenario, potential confounders include sexual history and sexual behavior. Users of the new device might differ in this regard from those using established IUDs. Information about sexual history and sexual behavior are typically not available from databases. Also, standard site based Non-Interventional Studies often have difficulties accessing such information. Thus, subject-based primary research approaches should be considered, specifically as valid long-term information is required to fulfill regulatory needs.

Our subject-based study example recruited a representative, real life (limited exclusion criteria) sample of women from routine private practice and public health care providers with wide geographical representation. Factors potentially associated with outcomes and exposures of interest were collected, including socio-demographics and lifestyle information. To ensure low attrition in a long term study, multiple-level follow-up measures were applied, using European or International Active Surveillance (EURAS/INAS) follow–up strategies (Fig. 2); for the supporting use of national registries to validate survival outcomes see also [3].

The main objective of a regulatory study is to collect valid product safety information. With a subject- based concept, patients are not limited by physician-defined adverse event categories: They can describe their complaints in their own words. As shown above, validating outcomes is important for subject based study quality. This is even more correct in a regulatory setting. With the EURAS/INAS concept [4], physicians validate events to ensure outcome validity: Interviewers and subjects complete follow-up questionnaires in a direct-to-subject contact. Treating physicians receive feedback on events for cross-checking and confirmation. During a second medical validation step, medical experts distinguish confirmed (definite / probable) events via blinded outcome adjudication.

Evaluating the effects of product exposure during pregnancy on pregnancy outcomes is a specific area of regulatory post marketing research. We need specific strategies to collect and validate pregnancy outcomes as well as infant outcomes, and to gather thorough information on exposure and potential confounders.
Database and site based approaches can rarely deliver thorough outcome, exposure, and confounding information. In contrast, subject-based 'Pregnancy Registries' collect exposure and confounding information directly from the pregnant mother. To ensure sample validity, a broad representation of women exposed to the product of interest during pregnancy can be recruited and retained beyond the duration of pregnancy. Combined with outcome validation by gynecologists and pediatricians as required, the subject-based pregnancy registry design thus derives valid information on pregnancy related medication risks and their effects (see *Pregnancy Registry case study box*).

Pregnancy Registry case study: Let's assume regulators need to understand the effects of medication taken during pregnancy on pregnancy outcomes and reasons for induced abortion. Information on reasons for abortion is not available from official sources and is typically not available from EMR or claims, but can be retrieved from the mother in subject-based approaches:

In the Ribavirin pregnancy registry [5], exposure of pregnant women to Ribavarin, as well as indirect exposure through the male sexual partner were assessed regarding pregnancy as well as infant outcomes.

Accessing valid infant outcomes provided by the pediatrician requires a link to the pediatrician. This link is hard to achieve by database linkage or by combining physician based information, but can be provided by the mother using a subject-based approach. This approach also permitted assessing reasons for induced abortion: In the Ribavirin pregnancy registry, induced abortion represented 22% of all evaluable pregnancy outcomes, and exposure to Ribavirin or potential birth defects were the main reasons given by mothers (49%) for induced abortion.

Subject Based Research Databases

Working with existing data or databases in secondary research can speed up the research process significantly and minimize cost. These are the main reasons why 'big data' is becoming more and more popular. Most health information databases rely on physicians providing access to their records (EMR) or on administrative data collected for reimbursement purposes (claims). For patient centered approaches, 'subject-based databases' would thus be of interest. Indeed, we see more and more analyses of unmonitored subject data from social networks and related sources, but outcome quality and epidemiological validity are often low, limiting the use of such data for regulatory or HTA purposes.

Figure 2: Typical follow-up procedures used to minimize attrition in long-term active surveillance [4].

Table 3: Subject Based Studies Summary

(Public) Health (Outcomes) Research & (Pharmaco-)Epidemiology					
DQM	Site Trial	Site Primary	Site Database	**Subject Primary**	**Subject Database**
	RCT…	OBS/NIS	EMR / Claims	PH / Survey	NHWS / plm ..
Sample / Epi Validity	? usually biased	(++) maybe biased	(++) depends on DB coverage	++	(++) depends on DB coverage
Independent variables / confounders	+ unless very personal	+ as available in routine	? as available in routine and in DB	? →++ support data improve quality	+ as available (subject report)
Admin Quality / Ethics / Regulatory	+ unless unethical	…requirements vary per research question, regulatory setting, internal requirements, country, study purpose etc…			
Attrition	? study design?	+ HCP dependent	? cross-sectional	++	? as available
Outcome (Clin/PRO/Eco)	++	++ as available in routine	+/- as available (PRO difficult)	? →++ support data → valid outcomes	+ as available (subject report)

Subject-based syndicated surveys provide an alternative: They deliver provider independent, unfiltered, health care information directly from the subject quickly and at reasonable cost. Other than with EMR or claims data, subjective information including health-related quality of life is often available, and survey management improves data and outcome validity compared to social network approaches. Main disadvantages relate to the fact that syndicated surveys are rarely designed for the research question in mind and thus may be limited with respect to the sample and confounders/outcomes available for analysis.

■ The National Health Interview Survey has been conducted by the US National Center for Health Statistics since 1957 and examines health status and health care access over 127,500 individuals.
■ The China Health and Nutrition Survey analyze the effects of health and nutrition on health outcomes covering 4,400 households and 26,000 individuals across 9 provinces.
■ PatientsLikeMe (www.patientslikeme.com) is a privately managed online network of 200,000+ voluntary participants and collects reports on treatment, conditions, and symptoms.
■ The National Health and Wellness Survey (NHWS, Kantar Health, www.kantarhealth.com) was initiated in 1988 and covers 10 countries and more than 250,000 subject reports today. Data are used for disease based health analytics as well as to assess market opportunities, estimate costs, gain insight into disease-specific segments, and to optimize value propositions and brand strategies.

Conclusions

Direct-to-subject studies, virtual trials, or subject-based research come close to optimally combining the most important quality factors in non-regulatory as well as regulatory health research, if performed with skill and operational excellence (see summary DQM in Table 3). Subject-based research allows for representative sampling of the population of interest. Engaging subjects in research can minimize attrition. Using research subjects as a link to different data sources ensures outcome validity.

A broad variety of topics can be covered by subject-based research, from pharmacoepidemiology and chemistry exposure [3] over occupational therapy and physiotherapy [6] to assessment of pharmaceuticals and devices. Direct-to-subject designs may also include opinion information, such as preference based approaches [7,8].

Whenever possible, study designs should include outcome validation to support evidence quality. This can be achieved if physicians adjudicate subject reported, or by means of official documentation such as maternity logs or public registries. In the area of regulatory post marketing requirement studies, authorities such as FDA and EMA actively recommend subject-based approaches. There definitely is value in subject-based research.

Acknowledgements

The author would like to thank P. Potthoff, S. Reed, S. Roberts and M. DiBonaventura for their contributions to this workshop..

References

[1] Hoeltz A, Hoeltz J, Potthoff P, et al.: Schwangerschaften und Geburten nach dem Reaktorunfall in Tschernobyl. Eine repräsentative Erhebung für die Bundesrepulik. Forschungsbericht. Available at: http://www.worldcat.org/title/schwangerschaften-und-geburten-nach-dem-reaktorunfall-in-tschernobyl-eine-reprasentative-erhebung-fur-die-bundesrepublik-deutschland-und-berlin-west-abschlubericht/oclc/165522854. [Accessed February 23, 2015]. [2] Roedler, Schwarz , Tsavarchidis et al.: Schwangerschaften und Geburten nach dem Reaktorunfall in Tschernobyl. Bundesamt für Strahlenschutz. Neuherberg 2011. [3] Potthoff P, Eichmann F, Klamert A. survival status in (pharmaco) epidemiological studies can be successfully investigate d using administrative residential registries. Value Health 2014;17:A544. [4] Dinger JC, Bardenheuer K, Assmann A. International active surveillance study of women taking oral contraceptives (INAS-OC Study). BMC Med Res Methodol 2009;9:77. [5] Roberts SS, Miller RK, Jones JK, et al. The ribavirin pregnancy registry: Findings after 5 years of enrollment, 2003-2009. Birth Defects Res A Clin Mol Teratol 2010;88:551-9. [6] Eichmann F, DiBonaventura M, Schoefer A. trends in physical and occupational therapy utilization in the US, and Western Europe. Value Health 2014;17:A406. [7] Pisa G, Eichmann F, Hupfer S. Assessing patient preferences in Heart Failure (HF) using conjoint methodology. Eur J Heart Fail 2014;16:354. [8] Pisa G, Eichmann F, Hupfer S. Assessing patient preferences in heart failure (HF) using conjoint methodology. Patient Pref Adhere 2015;9:1233-41. ■

Additional information:
The preceding article was based on the workshop, "Generating Evidence for Pharmacoepidemiology, Health Outcomes and Epidemiology through Direct-To-Subject Study Approaches" at the ISPOR 17th Annual European Congress, 8-12, 2014, Amsterdam, The Netherlands.

To view this presentation, go to: http://www.ispor.org/Event/ReleasedPresentations/2014Amsterdam#workshoppresentations.