PharmaSUG 2023 - Paper SI-05

What's the Story in Your Subgroup Analysis

Lucy Dai, AbbVie Inc, Abbott Park, IL;

ABSTRACT

Subgroup analysis is typically pre-defined in study analysis plan or ad hoc requested by study team or regulatory agency to further understand the treatment effect in the subgroups of interest. Through three subgroup analysis examples from real clinical trials this paper will try to give readers a glimpse of the impact they can have over the course of drug development and drug approvability.

INTRODUCTION

Pre-defined or ad hoc subgroup analysis in a study is often required to confirm the efficacy and safety consistency across different subgroups, for example analysis by age, by gender, by race, by center.

Common reasons of subgroup analysis are the following:

- To address concerns in some specific subgroups
- To explore whether the test drug is more efficacious or more harmful in a subset
- To provide supportive evidence to main findings
- To generate new hypothesis of drug effect
- To address regulatory queries

EXAMPLE I – GAME CHANGER

About thirty years ago a statistician on a rheumatoid arthritis (RA) clinical trial was asked by the study MD to do a subgroup analysis by stratification of prior medication use. No one expected that this analysis result changed the test drug development direction.

It was back in the 1980's, a test drug was on Phase 2 clinical trial to treat rheumatoid arthritis (RA) and osteoarthritis (OA). At that time there were two classes of RA drugs on the market, 1) Non-Steroidal Anti-Inflammatory Drugs (NSAID), 2) Disease Modifying Anti-Rheumatic Drugs (DMARD). These two classes are quite different in working mechanism and development approach. NSAID's are quick acting, symptom control, and more on pain relief, DMARD's on the other hand are slow acting (several months) and help slow down the disease progression.

Before this subgroup analysis the test drug was developed as NSAID class drug. The phase 2 study protocol was to compare test drug with Naproxen (an NSAID) on RA patients, it was a 5-year study with primary endpoint at 6 months and randomized with stratification on prior DMARD usage. Table 1 gives a quick summary of the study. Figure 1 is the ad hoc analysis requested by study clinical team, it is a subgroup analysis by randomization stratification of prior DMARD use [1].

Study team held a discussion meeting, and it was obvious test drug has better efficacy than Naproxen (an NSAID), one of the attendees raised the point that result from this subgroup analysis indicates test drug may be a DMARD, instead of NSAID. This is a game changer, because DMARD has a totally different working mechanism, marketing strategy and drug approvability. Sponsor's further development demonstrated that test drug has DMARD efficacy. Surprisingly this subgroup finding served as the change driver of pipeline development direction.

Table 1.

Time	1980's
Study phase	Phase 2
Target indication	Rheumatoid arthritis (RA) and Osteoarthritis (OA)
Study design	Randomized, parallel controlled
Study length	5 year
Primary analysis	At 6 months
Primary endpoint	Change of physician assessment of disease activity at month 6 from baseline: No symptom (1) Mild (2) Moderate (3) Severe (4) Very severe (5)
Primary endpoint result	Test drug -0.76, Naproxen -0.57 p=0.0103

Table 1. Study Basic Information

Figure 1.

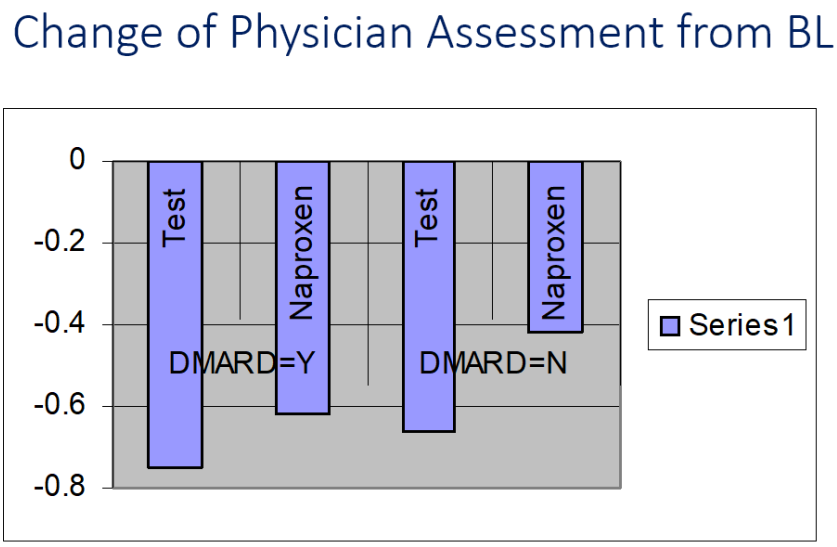

Figure 1. Primary Endpoint Subgroup Analysis by Prior DMARD Usage [1]

EXAMPLE II – HYPOTHESIS GENERATING VS. HYPOTHESIS TESTING

In 1990's there was a clinical trial PRAISE-1 (Prospective Randomized Amlodipine Survival Evaluation-1 study), it randomly assigned 1153 patients with severe chronic heart failure and ejection fractions of less than 30 percent to double-blind treatment with either placebo (582 patients) or amlodipine (571 patients) for 6 to 33 months, while their usual therapy was continued. The randomization was stratified on the basis of whether patients had ischemic or nonischemic causes of heart failure [2].

After a median follow-up of 13.8 months, the incidence of the primary endpoint of all-cause mortality or cardiovascular hospitalization was similar in both groups (amlodipine: 222 of 571 [39%]; placebo: 246 of 582 [42%]; hazards ratio [HR] [95% confidence interval] (CI): 0.91 [0.76 to 1.10]). This neutral effect on the primary composite endpoint was also reported on the most important secondary endpoint: death from all causes [3].

However, further probing of the data exploring multiple subgroups led to a potentially important observation: amlodipine seemed to reduce the risk of experiencing the primary composite endpoint as well as almost cutting in half the risk of death in those classified in the nonischemic stratum, p value for interaction = 0.004 [3].

At this point, the intriguing subgroup analysis finding generated a new hypothesis that amlodipine could benefit patients with heart failure due to a nonischemic cardiomyopathy. Soon after a second large-scale trial PRAISE-2 was specifically designed to test the hypothesis observed in PRAISE-1. But to everyone's disappointment PRAISE-2 results were not statistically significant. Even when patients with nonischemic cardiomyopathy in both PRAISE-1 and PRAISE-2 were combined, there was no evidence of a favorable effect of amlodipine on mortality.

As discussed in the paper by Milton Packer [3], the baseline characteristics of nonischemic patients in the two trials were similar, and the trials were carried out using virtually identical protocols and similar investigators. Furthermore the paper stated that the totality of available evidence suggests that the benefits of amlodipine seen in PRAISE-1 were related to chance, the encouraging findings in PRAISE-1 were based on a subgroup analysis of a secondary endpoint in a trial that failed to achieve its primary endpoint.

Table 2. Effect of Amlodipine on Survival [3]

Table 2 **Effects of Amlodipine on Mortality in Patients With Severe Heart Failure Due to a Nonischemic Cardiomyopathy**

Study	Amlodipine	Placebo	HR (95% CI)	p Value†
PRAISE-1 (9)	45/209	74/212	0.54 (0.37–0.78)	0.001
PRAISE-2	278/827	262/827	1.09 (0.92–1.29)	0.3
PRAISE-1 (9) + PRAISE-2‡	322/1036	332/1039	0.97 (0.83–1.13)	0.6

EXAMPLE III – PRODUCT PACKAGE INSERT

This is an example from a safety subgroup analysis by age. Figure 2 below is a screenshot from a FDA approved package insert. It has different dosage recommendations to adults 65 years of age and older, while adults less than 65 years of age can increase to 30mg daily if an adequate response is not

achieved. How was this determined? It's driven by risk-benefit evaluated from pre-defined subgroup analysis by age result.

Figure 2.

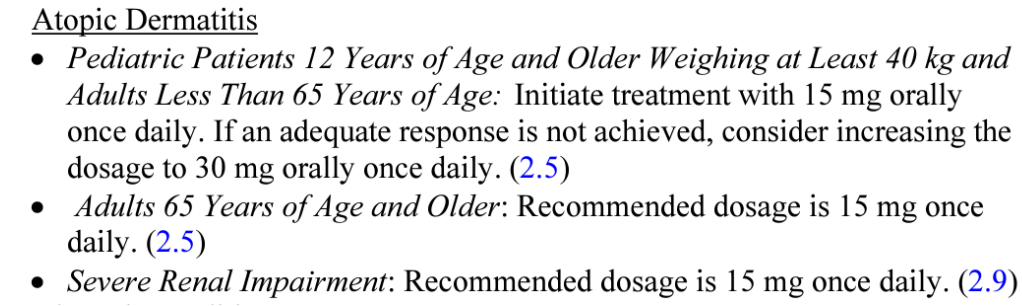

Figure 2. FDA approved label in 2022-10-21 [5]

In Geriatric section from FDA approved package insert, it states "*Of the 2583 patients treated in the three Phase 3 clinical trials, a total of 120 patients with atopic dermatitis were 65 years of age or older, including 6 patients 75 years of age. No differences in effectiveness were observed between these patients and younger patients; however, there was a higher rate of serious infections and malignancies in those patients 65 years of age or older in the 30 mg dosing group in the long-term trials.*" [5]

The ISS table [5] that can be shared in this paper is at the below that shows the ISS analysis results by week 16, it has 115 patients that were 65 years of age or older, serious infection and malignancies rates were higher among patients 65 years of age or older who received 30 mg dosing.

TABLE E4. Adverse events by age group in the All Upadacitinib Exposure analysis

	<65 years		≥65 and ≤75 years	
Events (E/100 PY) [95% CI]	**UPA 15 mg N = 1191 PY = 1326.3**	**UPA 30 mg N = 1179 PY = 1341.1**	**UPA 15 mg N = 48 PY = 47.1**	**UPA 30 mg N = 67 PY = 73.1**
Adverse events of special interest				
Serious infection	32 (2.4) [1.7, 3.4]	33 (2.5) [1.7, 3.5]	0	6 (8.2) [3.0, 17.9]
Opportunistic infection (excluding tuberculosis and herpes zoster)†	22 (1.7) [1.0, 2.5]	26 (1.9) [1.3, 2.8]	0	1 (1.4) [0.0, 7.6]
Herpes zoster	47 (3.5) [2.6, 4.7]	72 (5.4) [4.2, 6.8]	1 (2.1) [0.1, 11.8]	2 (2.7) [0.3, 9.9]
Active tuberculosis	1 (<0.1) [0.0, 0.4]	1 (<0.1) [0.0, 0.4]	0	0
NMSC†	PY = 1322.4 4 (0.3) [0.1, 0.8]	PY = 1339.1 4 (0.3) [0.1, 0.8]	PY = 47.1 0	PY = 71.8 1 (1.4) [0.0, 7.8]
Malignant neoplasms excluding NMSC†	PY = 1326.1 2 (0.2) [0.0, 0.5]	PY = 1340.8 4 (0.3) [0.1, 0.8]	PY = 47.1 0	PY = 72.9 3 (4.1) [0.8, 12.0]

CONCLUSION

This paper intents to give readers a peek into the effect that subgroup analysis can exert over the course of drug development through these three real life examples. Subgroup analysis could be more than just to provide supportive evidence to main findings. They could push study team to reconsider the working mechanism of a test drug and change the development direction, they could generate new hypothesis to test if investigational drug can potentially benefit a specific subgroup of patients, they could also make their way to the final commercial label.

Readers need to be reminded of a statement from ICH E-9 [6], "When exploratory, these analyses should be interpreted cautiously; any conclusion of treatment efficacy (or lack thereof) or safety based solely on exploratory subgroup analyses are unlikely to be accepted."

REFERENCES

[1] Subgroup analysis class notes by Naitee Ting
https://higherlogicdownload.s3.amazonaws.com/AMSTAT/87de5f9b-d337-4398-8c8e-a929aaace48b/UploadedImages/Mini_Conference/2020/TingPresentation.pdf

[2] Marc A. Pfeffer, MD, PHD, Hicham Skali, MD, MSC "PRAISE (Prospective Randomized Amlodipine Survival Evaluation) and Criticism" JACC: Heart Failure Vol. 1, No. 4, 2013
https://www.sciencedirect.com/science/article/pii/S2213177913001923?via%3Dihub

[3] Milton Packer MD, Peter Carson MD, Uri Elkayam MD, Marvin. Konstam MD, Gordon Moe MD, Christopher O'Connor MD, Jean-Lucien Rouleau MD, Douglas Schocken MD, Susan A. Anderson MS, David L. DeMets PhD, for the PRAISE-2 Study Group. "Effect of Amlodipine on the Survival of Patients With Severe Chronic Heart Failure Due to a Nonischemic Cardiomyopathy" JACC: Heart Failure Vol. 1, No. 4, 2013
https://pubmed.ncbi.nlm.nih.gov/24621933/

[4] FDA approved Rinvoq label
https://www.accessdata.fda.gov/drugsatfda_docs/label/2022/213871s000lbl.pdf

[5] Emma Guttman-Yassky MD, PhD, Jacob P. Thyssen MD, PhD, DmSci, Jonathan I. Silverberg MD, PhD, MPH, Kim A. Papp MD, PhD, Amy S. Paller MD, Stephan Weidinger MD, H. Chih-ho Hong MD, Barbara Hendrickson MD, Deanne Dilley MPH, Allan R. Tenorio MD, Barry Ladizinski MD, MPH, MBA, Alvina D. Chu MD, John Liu MD, MS, Alan D. Irvine MD, DSc "Safety of upadacitinib in moderate-to-severe atopic dermatitis: An integrated analysis of phase 3 studies" J ALLERGY CLIN IMMUNOL VOLUME 151, NUMBER 1
https://www.sciencedirect.com/science/article/pii/S0091674922013276

[6] ICH E9 statistical principles for clinical trials - Scientific guideline | European Medicines Agency (europa.eu)

This manuscript was sponsored by AbbVie. AbbVie contributed to the design, research, and interpretation of data, writing, reviewing, and approved the content. All authors are employees of AbbVie Inc. and may own AbbVie stock.

ACKNOWLEDGMENTS

The author wishes to thank Naitee Ting for aspiring me to present subgroup topic and allowing me to borrow his class slides. The author also appreciates the thoughtful review comments from Lin Bai, Holly Peterson and the generous help from Abbvie Immunology group statisticians Yuanyuan Duan, Yihan Li, Xiaofei Hu and Yang Yang, Immunology group clinical team member John Liu.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Please contact the author at:

Lucy Dai
Work Email: lucy.dai@abbvie.com
Work Phone: 832-263-4819