

Indirect treatment comparisons and biologics



David Mauger, PhD,^a and Andrea J. Apter, MD, MSc, MA^b Hershey and Philadelphia, Pa

Key words: Asthma, biologic, monoclonal antibody, benralizumab, reslizumab, systematic analysis

There has been much recent excitement among both patients and clinicians over the development and approval of new biologic medications, mAbs, for asthma, the first new class of medications targeted at pathobiological pathways since antileukotrienes became available in the 1990s.^{1,2} Clinicians want to know how to use these medications and which of those with similar mechanisms are likely to be most effective for their patients.³ These medications are very expensive (Table I), and their duration of use, long-term benefits, and adverse effects are unknown. In addition, biologics with similar mechanisms, such as blocking the actions of IL-5, have not been directly compared for safety or efficacy in head-to-head clinical trials. However, 2 recent publications have constructed indirect comparisons based on separate placebo-controlled clinical trials.^{4,5} Interestingly, the results are not completely congruent.

An indirect comparison estimates the difference between 2 treatments, say X and Y, through a common comparator. Registration trials for all new drugs include placebo, which makes it the logical choice for a common comparator. Thus the difference between the treatments can be represented as follows:

$$T_X - T_Y = [T_X - P_X] - [T_Y - P_Y] - [T_X - P_Y],$$

where *T* denotes treatment and *P* denotes placebo. Under this framework, $[T_X - P_X]$ and $[T_Y - P_Y]$ can be estimated from existing placebo-controlled trials for the 2 treatments. For obvious reasons, however, trials comparing P_X and P_Y will never be done, which necessitates an assumption that the difference between P_X and P_Y (whatever it is) is small relative to $[T_X - P_X]$ and $[T_Y - P_Y]$.

However, even if it is reasonable to assume that the difference between P_X and P_Y is very small, there is a more fundamental

condition that must also be met for the indirect comparison to be valid. The justification for the indirect comparison rests on the assumption that the results of a hypothetical head-to-head trial comparing treatments X and Y can be approximated by using results from separate trials of X and Y with a common comparator. This would be reasonable if these separate trials were done under the same conditions and in the same patient populations.

A network meta-analysis (NMA) builds evidence by combining results from head-to-head trials (direct evidence) with results from common comparator trials (indirect evidence), whereas a traditional meta-analysis includes only head-to-head trials.⁶ Although some degree of variation across study populations is acceptable in a traditional meta-analysis, it can be disastrous in an NMA. Because an NMA combines 2 different types of trials, it necessitates the assumption of “consistency,” which means that the indirect and direct evidence is congruent; that is, if the indirect evidence from the common comparator trials included in the NMA indicates that treatment X is superior to treatment Y, then such differences also should be seen in head-to-head trials. Unfortunately, these recent NMAs do not include any head-to-head trials because none have been done. Therefore although NMA is a statistically sound methodology that has been applied in many settings, there is significant vulnerability in these recent publications because it is not possible to objectively assess the consistency assumption, and we are left to evaluate it subjectively. Establishing that the trials were done under equivalent conditions would go a long way in this regard.

PICOS (Population, Intervention, Comparator, Outcomes, Study Design) is an ideal framework for this assessment (Table II). With respect to the patient population aspect, it is not absolutely necessary for the trials included in the NMA to be identical in all characteristics, but it is critical that they be very similar in all characteristics that are important for trials of biologics in asthmatic patients because those are the ones most likely to affect consistency. How does one determine these characteristics? An obvious place to start is inclusion criteria and suspected predictors of asthma outcomes and treatment responses. Both of the recent NMAs address this issue by comparing the baseline characteristics of the patient populations included in their analyses. Busse et al⁵ take the step of formally evaluating which of these characteristics are likely to be effect modifiers for biologics. For example, if studies differ by the inclusion of patients with different blood eosinophil counts (≥ 0 vs ≥ 350 cells/ μ L), their response to the drug could be different. However, the bar they set is so high (ie, “conclusive evidence of effect modification”) that they could be overlooking important suspected modifiers that have not been proved yet. In any case, if the patient populations in the trials included in an NMA differ in terms of important characteristics, then we have reason to be concerned about the consistency assumption. One approach, which is applied in both reports, is to reduce this concern by extracting subgroups of patients that are similar across all of the trials included. Although this does help increase our confidence in

From the ^athe Department of Public Health Sciences, Penn State College of Medicine, Hershey, and ^bthe Division of Pulmonary, Allergy Critical Care Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia.

Reprinted in J Allergy Clin Immunol Pract 2019;7:131-3.

Disclosure of potential conflict of interest: D. Mauger has received grants from the National Heart, Lung, and Blood Institute (NHLBI), serves on Data Monitoring Committees for Biocryst and Novartis, and his institution has received in-kind donations in the form of drug products from Boehringer Ingelheim, GlaxoSmithKline, Merck, and TEVA in support of clinical trials funded by the NHLBI, and has received financial donations from AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, Genentech, Sanofi, and TEVA in support of a clinical research study funded by the NHLBI. A. J. Apter has received grants from the National Institutes of Health/NHLBI and Patient-Centered Outcomes Research Institute. She is an associate editor of the *Journal of Allergy and Clinical Immunology* and is a consultant for UpToDate.

Received for publication October 29, 2018; revised November 8, 2018; accepted for publication November 9, 2018.

Corresponding author: Andrea J. Apter, MD, MSc, MA, 829 Gates/HUP, 3600 Spruce St, Philadelphia, PA 19104. E-mail: andrea.apter@uphs.upenn.edu.

J Allergy Clin Immunol 2019;143:84-6.
0091-6749/\$36.00

© 2018 American Academy of Allergy, Asthma & Immunology
<https://doi.org/10.1016/j.jaci.2018.11.005>

TABLE I. Average wholesale price* of one dose and for a 52-week year of mepolizumab, reslizumab, and benralizumab for a 70-kg adult

Biologic	Dose	Cost of 1 dose for a 70-kg subject	Cost for a 52-wk year
Mepolizumab	100 mg SQ q4W	\$3,442.40	\$44,751.20
Reslizumab	3 mg/kg IV q 4	\$107.52/mL (~\$2150)	\$27,950.00
Benralizumab	30 mg SQ q4W for 3 doses and then q8W	\$5,702.53	48,471.51, year 1; \$42,768.98, thereafter

IV, Intravenous; Q, every; SQ, subcutaneous; W, week.

*The average wholesale price (AWP) is the average value at which wholesalers sell drugs to physicians, pharmacies, and other customers. The AWP is a generally accepted standard measure for calculating the cost of a medication. It is typically 20% greater than the value for which a manufacturer sells their products to distributors and large customers.

TABLE II. The PICOS framework, outlined in this table, is used to assess whether the component studies in meta-analyses are sufficiently similar. This table includes examples of elements used in the component studies that must be compared

Element		
P	Population	Age Sex Criteria used to define severe eosinophilic asthma Baseline eosinophil counts Baseline asthma control Baseline asthma-related quality of life Baseline FEV ₁ percent predicted Baseline medications, including dose of inhaled corticosteroid Baseline asthma severity (GINA level) Number of exacerbations in year before enrollment Smoking history
I	Intervention	Anti-IL-5 biologic Duration of treatment
C	Comparator	Placebo Other asthma medications Comorbidities
O	Outcomes	Clinically significant exacerbations Asthma control Asthma-related quality of life Emergency department visits/hospitalizations
S	Study type	Randomized controlled trial

GINA, Global Initiative for Asthma.

the consistency assumption, it narrows the generalizability of the comparison to a smaller population.

With respect to the intervention and comparator aspects of PICOS, there is also reason for concern about these NMAs. For example, both NMAs include trials that differ by duration of intervention therapy, ranging from 15 to 56 weeks.^{4,5} In addition, although all of the trials included in these NMAs used placebos, that does not necessarily make them equivalent comparators. Because asthma outcomes are affected by background controller medications (eg, inhaled corticosteroids/long-acting β -agonists), we cannot be confident that the placebo arms of the component trials offer comparable results.

In addition to patient characteristics, interventions, and comparators, compatibility of outcomes must also be examined.⁷ For these NMAs, they concern lung function, asthma

control, and exacerbations. Lung function measurement is highly standardized and thus an ideal outcome. With respect to asthma control, the 2 most commonly used tools are the Asthma Control Questionnaire, which queries symptoms over the past week,⁸ and the Asthma Control Test,⁹ which asks for symptoms over the past 4 weeks. In these NMA studies the Asthma Control Questionnaire was uniformly used. Exacerbations, on the other hand, were defined less uniformly across the trials included in these NMAs. Exacerbations are an inherently difficult outcome to standardize because they entail subject assessments at points. First, the patient must make the decision to seek medical care, and second, a physician who might or might not be familiar with the study protocol must make the decision to initiate systemic corticosteroids. Most clinical trials try to address this by supplying patient asthma action plans and specifying criteria for initiating systemic corticosteroids, but significant variation remains.

With respect to the study design component of PICOS, the fact that patients in both the placebo and active arms improve suggests another mechanism of benefit: that these drugs, which must be given under medical supervision, are a form of directly observed therapy. This is because patients are seen regularly by medical personnel to receive the study medication, any illness is likely to be recognized more quickly, and treatment is likely to be initiated more promptly. With all of the differences in the trials available for inclusion, it is not surprising that these 2 NMAs comparing mAbs report different results.^{4,5}

As we try to learn about and compare these new biologics, it is critical to understand the limitations of the NMA. Most important, they must include some head-to-head trials making direct comparisons to rigorously and objectively evaluate the critical assumption of consistency. In addition, different design choices can lead to inclusion of different patients and result in different outcomes that might be valid for the cohort studied but not provide information on patients with different characteristics. Those design choices (PICOS) should be detailed in research reports.

It is also important not to lose sight of the fact that there are other aspects of caring for patients with asthma that NMAs cannot consider. For example, they do not provide information on patients' and clinicians' goals, priorities, and beliefs. Initial studies on new medications tend to focus on patients most likely to benefit from the medication and might not include patients with comorbidities, such as hypertension, diabetes, or obesity or exposures to tobacco, pollutants, and other toxic agents. These considerations tell us that clinicians and patients must wait until more patients are exposed to these new medications for longer

periods and under real-life conditions to fully appreciate their benefits and adverse events. Finally, problems of cost must be overcome.

REFERENCES

1. Drazen JM, Israel E, O'Byrne PM. Treatment of asthma with drugs modifying the leukotriene pathway. *New Engl J Med* 1999;340:197-206.
2. Chu EK, Drazen JM. Asthma: one hundred years of treatment and onward. *Am J Respir Crit Care Med* 2005;171:1202-8.
3. Manka LA, Wechsler ME. Selecting the right biologic for your patients with severe asthma. *Ann Allergy Asthma Immunol* 2018;121:406-13.
4. Casale TB, Pacou M, Mesana L, Farge G, Sun SX, Castro M. Reslizumab compared with benralizumab in patients with eosinophilic asthma: a systematic literature review and network meta-analysis. *J Allergy Clin Immunol Pract* 2019; 7:122-30.
5. Busse W, Chupp G, Nagase H, Albers FC, Doyle S, Shen Q, et al. Anti-IL-5 treatments in patients with severe asthma by blood eosinophil thresholds: indirect treatment comparison. *J Allergy Clin Immunol* 2019;143:190-200.
6. Rouse B, Chaimani A, Li T. Network meta-analysis: an introduction for clinicians. *Intern Emerg Med* 2017;12:103-11.
7. Busse WW, Morgan WJ, Taggart V, Togias A. Asthma outcomes workshop: overview. *J Allergy Clin Immunol* 2012;129(suppl):S1-8.
8. Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. *Eur Respir J* 1999; 14:902-7.
9. Nathan RA, Sorkness CA, Kosinski M, Schatz M, Li JT, Marcus P, et al. Development of the asthma control test: a survey for assessing asthma control. *J Allergy Clin Immunol* 2004;113:59-65.