Jason Spangler, MD, MPH, FACPM
Executive Director,
Value, Quality, and Medical Policy

U.S. Health Policy & Reimbursement
601 Thirteenth Street, NW
Twelfth Floor
Washington, DC 20005
Phone: 202.585.9659
Fax: 202.585.9730
Email: jspangle@amgen.com
www.amgen.com

July 16, 2018

VIA ELECTRONIC DELIVERY

Tamara Syrek Jensen, JD
Director, Coverage and Analysis Group
Centers for Medicare and Medicaid Services
7500 Security Boulevard
Baltimore, MD 21244

Re: Medicare Program; Meeting of the Medicare Evidence Development and Coverage Advisory Committee - August 22, 2018

Dear Ms. Syrek Jensen,

Amgen Inc. (Amgen) appreciates the opportunity to comment on the voting questions for the Medicare Evidence Development and Coverage Advisory Committee (MEDCAC) Meeting on Chimeric Antigen Receptor (CAR) T-Cell Therapy and Patient Reported Outcomes (PROs). Amgen strongly supports ensuring patient-centricity in developing innovative therapies, and we have research efforts underway with clinical collaborators to incorporate patient-centric outcomes in our clinical research programs. In advance of the MEDCAC meeting, we submit the below recommendations and considerations on use of PROs and expansion of the field of patient-derived data collection for years to come.

PROs have tremendous promise considering the explosion in data streams becoming available for concurrent study. Determining how PRO tools should be incorporated into clinical studies is a critically important goal. We look forward to the MEDCAC meeting and expect it to be a helpful discussion on ways to improve PROs and incorporate them into the coverage process. We agree with CMS's goal of seeking recommendations from the MEDCAC Panel regarding how existing PRO assessment tools should be incorporated into future clinical studies, including future clinical studies on CAR T-cell therapy, and with the proposal that the MEDCAC focus on important characteristics of a PRO assessment tool. We broadly agree with CMS's proposal for the MEDCAC panel to assess whether scientific evidence supports various types of outcome assessments, study design characteristics, study duration, and suitable controls for applying PROs to health outcomes research. We likewise agree that it is important for the MEDCAC to

explore the challenges regarding the validity, reliability, and generalizability of PRO assessments and ways to improve those assessments.

We are concerned, however, that focusing on existing "specific PRO assessment tools" could engender a narrow instrument (not patient) focused approach that may not be appropriate to achieving the agency's objectives. Patient experience data plays a significant role in helping to identify unmet medical needs and important clinical outcomes that have yet to be studied, thereby informing future PRO development and selection, as well as analyses and communication of benefit-risk. We encourage the agency to ensure that questions regarding the number of instruments, study design, and study length are not so narrow as to discourage efforts to expand collection of patient-derived data in multiple settings, using multiple additive methods, for multiple purposes. To accomplish this, Amgen respectfully recommends a broader approach to the application of PROs in the evolving oncology setting, involving expanding and/or reframing the current questions to match the broader set of issues facing the use and interpretation of PROs in the clinical setting.

Background

In the U.S., the term "Patient Reported Outcome" has become an accepted umbrella term for information collected directly from the patient, generally using structured and standardized questionnaires developed utilizing accepted psychometric principles that produce data without the involvement of a health professional. This approach is distinct from data obtained by observation or history-taking by a physician, obtained from laboratory or imaging, or apparent as a physical sign of a disease. In practice, of course, physicians and other health professionals are collecting a continuous stream of real-time data as reported by the patient. Pain, activity level, decline or improvement in reported function, new symptoms, etc. are all areas of routine inquiry in disease evaluation and management. The distinguishing feature of PROs is therefore the consistent, standardized, psychometrically-validated approach to collecting information that may be overlooked or not systematically captured by traditional clinical management.

PROs are increasingly being developed as additional clinically-relevant outcomes in clinical trials. Development of instruments for the clinical trial setting is one of the most rigorous applications of PROs and the application that has typically received the most attention as new instruments and analytic techniques are developed and validated. Regulatory agencies have issued various guidance on the steps required to obtain consideration of PRO data for product labeling.

Over time, PRO data take on more and more meaning and are subject to better interpretation as data from hundreds or thousands of other patients who have answered the same exact questionnaire become available for concurrent analysis. In general, the more widely studied, used, and characterized an instrument becomes in multiple settings and situations, the more it is said to be "validated". This means of course, that the "validity" of a PRO is a continuous construct that is improving with each new data point obtained for the PRO. Increasing validity

is desirable because it implies more ability to extrapolate results to populations that may vary from previously studied populations, and to successfully use the instrument in more heterogeneous settings. It does not necessary mean that the instrument has been studied in the same exact settings as one under consideration.

The torrent of digitized medical data becoming more widely available promises to greatly augment the ability to both collect and interpret data from PROs. We are encouraged by new opportunities to broaden the evidence base for PROs with the advent of multiple new technologies in health care such as electronic medical records (EMRs), mobile technology, activity trackers, and real time biometric data collection. Real-time PRO data will almost certainly be critical to supplementing and interpreting these "hard" data streams, while advancing PRO development and validation for use in far more settings

MEDCAC Discussion Should Reflect Broad Approach to PRO Value and Generalizability

As a global matter, we believe that it is important for CMS's questions and the MEDCAC's analysis to recognize the need for a rigorous but workable approach to the value and generalizability of PRO instruments. For instance, we are concerned that CMS's question about "validation", which is the first MEDCAC question, poses the risk of an overly narrow view of generalizability that could result in MEDCAC missing the opportunity to encourage and benefit from the widespread application of high-quality instruments in cancer trials that routinely enroll Medicare subjects. These data play an important role when combined with increasing amounts of companion data for supplementary analyses and control, to supplement interpretation of clinical endpoints that are often sterile and devoid of patient-derived context.

As context, it is useful to note that much of the discussion about PROs in clinical trials has been narrowly focused on registration of PRO endpoints and/or obtaining product labeling incorporating PRO data. The PRO Guidance from the Food and Drug Administration (FDA) is highly prescriptive and designed primarily to maximize the internal validity of *a priori* selected PRO endpoints for specific populations addressed in specific trial settings.[1,2] This is not an unreasonable goal for the FDA and manufacturers during registration discussions since the statutory standard of "two adequate and well-controlled clinical trials" is a starting point for what information may be included in a product label and used in promotional discussions. FDA and manufacturers share a common interest in crafting product labeling that is rigorous, clear

---

[1] US Department of Health and Human Services. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. 2009.

[2] US Department of Health and Human Services. Patient Preference Information—Voluntary Submission, Review in Premarket Approval Applications, Humanitarian Device Exemption Applications, and De Novo Requests, and Inclusion in Decision Summaries and Device Labeling 2016.

and unambiguous, and applicable to the use of the drug in diverse clinical settings. Indeed, many subjects participating in oncology clinical trials are 65 years of age or older and covered by Medicare.

It is important to note in this regard that opportunities for use of PRO data for additional decision making in multiple treatment settings will almost always exceed FDA's capacity and willingness to recognize such effects in labeling, or the instrument developer/user's ability to anticipate, analyze and publish on hundreds of potential scenarios. The converse is also important: PRO effects recognized in labeling should be afforded substantial credence given the rigor of this setting, and thus may be more widespread and generalizable than labeling addresses because of the rigor by which they were originally captured and scrutinized.

Outside of the labeling and clinical trial context, investigators and clinicians agree that additional data from supplementary, good quality PRO instruments can be very useful in concert with traditional endpoints, and certainly when combined with new types of periodic or real-time biometric data. Importantly, a nuanced view of the generalizability of such instruments must be taken into account when assessing their reliability in the Medicare coverage context. Because increasing validation of any instrument is a consequence of widespread use and analysis, many rigorous instruments have not been specifically validated for use in multiple and specific populations such as Medicare beneficiaries, meaning that the typical question of whether a PRO instrument is validated in a given setting would often be answered in the negative. When one further considers natural variation in past study designs regarding patient diagnosis, clinical status, etc., it is even less likely that the instrument in question has been used in the "same" settings. Sometimes the standard of "exact" setting is also applied, which is an impossibly high bar for any instrument to clear. PRO instruments considered in a regulatory setting are therefore often held to a higher validity standard than many accepted clinical endpoints.

Moreover, clinical insights are generally obtained through multiple measures of a patient's condition looked at together rather than in isolation. The same is true for PRO instrumentation. In general, PRO instruments are: 1) general quality of life, well-being, or global status assessments; 2) disease-specific questionnaires designed to capture more granular concerns that might be missed with general instruments; or 3) symptom- or sign-specific measures that are designed for even deeper focus on a specific area of clinical performance. It may be appropriate to use multiple instruments to ascertain the patient's condition, and well-designed PRO programs often do just that.

We believe that these nuances should be taken into account in reviewing the validity and generalizability of PROs. We are thus concerned that the line of MEDCAC questioning about individual instruments may result in inadvertently pessimistic conclusions about the generalizability of a broad-based platform of PRO data to the Medicare population, and for use in future oncology studies that are likely to be generalizable to the Medicare and other populations.

Specific Comments on Questions and Recommendations

Question 1

How confident are you that each of the following PRO assessments are valid and generalizable to the Medicare population?

Amgen believes the way the question is posed may preempt real opportunities to expand the use of high quality instruments in more settings, with more patients. Specifically, in the absence of data to the contrary, every one of the PRO questionnaires listed in Question 1 and additional questionnaires could be valid and generalizable specifically to the Medicare population. Data may, in fact, already exist to support this. Given the observation above regarding the rigor of the regulatory setting, it is highly likely that PRO effects observed in well-designed clinical studies are much more generalizable than suggested by labeling or the absence of labeling. Further, these instruments can be made to be more valid in *any* population with additional studies, companion data or supplementary sub-studies. Instrument performance by insurance coverage has not typically been a primary analytic question for instrument development, though arguably it should be, given the importance of insurance coverage in care delivery.

Developing PRO questionnaires, measuring PRO in clinical trials, and demonstrating the validity of the PRO questionnaires have formed a scientific discipline that started multiple decades ago involving behavioral scientists, psychometricians, health services researchers, and others. For more than a decade, regulatory agencies, particularly FDA and European Medicines Agency (EMA), have acknowledged the value of directly collecting patients' experiences via PRO questionnaires by issuing PRO guidance, reflection papers (EMA Reflection paper for HRQL measures, 2005), and a framework for PRO data collection in cancer clinical trials.[3,4,5]

When PRO questionnaires are developed, patient focus groups, patient interviews, and/or clinician interviews are usually assembled to demonstrate the content validity of the questionnaires. In addition, psychometric validations are usually conducted in a broad-based group of patients to ensure that the questionnaires are reliable, generate repeatable results, and can detect a clinically meaningful change often in a variety of settings. Given the rigorous design of most of the PRO instruments being considered in this question, it is thus highly likely that results will be resilient to settings of use, and that meaningful data will be obtained and directionally similar conclusions reached from analysis of these instruments specifically for

---

[3] US Department of Health and Human Services. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. 2009.

[4] Reflection Paper on the Regulatory Guidance for the Use of Health-Related Quality of Life (HRQL) Measures in the Evaluation of Medical Products (July 2005).

[5] Kluetz PG, et al. Focusing on Core Patient-Reported Outcomes in Cancer Clinical Trials: Symptomatic Adverse Events, Physical Function, and Disease-Related Symptoms. Clin Cancer Res;22(7):1553–8.

Medicare patients in cancer, despite lack of formal, published validation in the Medicare population alone.

Besides the seven questionnaires listed in Question 1, other PRO questionnaires could also be valid and generalizable to Medicare patients in cancer, including, but not limited to, non-disease-specific instruments EQ5D[6], 36-item short form survey, (SF36), health utilities index (HUI); cancer-generic instruments functional assessment of cancer therapy (FACT); cancer type specific instruments BREAST-Q[7], and other potentially newly developed questionnaires that follow the professional societies' recommended best practices.

Question 2

Considering all PRO assessments in question 1 with greater than or equal to score 2.5, please vote whether or not those PRO assessments combined have available supporting evidence on each of the following desired characteristics.

We are concerned that Question 2 poses similar risks to Question 1, as it is phrased in a restrictive manner that may invoke the regulatory validity standard in all potential decisions and conclusions and may preempt clinically appropriate decision making and the important work that should be encouraged by panels like the MEDCAC.

As discussed in the response to Question 1, PRO questionnaires developed using methods tested and refined for decades are intended to be generalizable over many situations, but not every situation can be anticipated and studied before the instrument is released for widespread use. It is likely that most of the instruments noted in Question 1 will perform well in most of the situations posed in Question 2. Ultimately, additional data collected in daily use is what makes researchers increasingly comfortable in interpreting every new result.

We recommend that CMS approach Questions 1 and 2 by grouping the instruments by General, Disease, and Symptom/Sign-specific, and then asking the MEDCAC "Which of the following instruments, when used in the Medicare population, would likely return results that would be interpretable and generalizable?" In addition to label and formal review data, Panel members could base their response on: A) professional opinion; B) breadth of data available from all populations previously studied and published; and C) cohort analysis (likely highly limited) available for Medicare subgroups. This would provide an opportunity to allow the experts to be future-focused, while also looking back at the sum of current available evidence in a variety of use settings. Another useful line of inquiry could be whether there are any data that would prevent or disqualify the instrument from being used, such as conflicting or contrary data from a more Medicare-like setting.

---

[6] https://euroqol.org/eq-5d-instruments/

[7] https://eprovide.mapi-trust.org/instruments/breast-q#basic_description

Question 3

How confident are you that each of the following assessment intervals are appropriate measurement periods for a valid PRO assessment?

- *Variable event-dependent frequency interval (i.e. upon admission and after discharge)*
- *Fixed time-dependent frequency interval (i.e. weekly, monthly, or yearly)*

PRO assessment schedules in clinical trials should be designed to increase the likelihood of accurately capturing the conditions or phenomena of interest, while avoiding posing an undue burden for the subject or study, and not imposing a Hawthorn effect on the patients themselves. Ultimately, this is a study design judgement. It is not uncommon in oncology clinical trials that both impacts of an important event (e.g., hospitalization, adverse events, relapses) and impacts of a treatment over the entire treatment time are of interest. For example, when performing analyses looking at the impact of important events on a patient's well-being, it is useful to have a "general background" of assessments from a theoretically unbiased and uniform random set of assessments. It is also important to have the proximate, point-in-time assessment that captures the impact of shorter duration events that might be missed if triggered by an event such as a side effect or admission to hospital.

Therefore, the answer to voting question 3 is that both choices could be appropriate.

Question 4

How confident are you that a PRO assessment over the course of the following study durations identifies a meaningful durable treatment effect with a valid PRO?

- *6 months*
- *12 months*
- *24 months*

We believe that Question 4 has implications that are much broader than planning oncology trials. This question forces the MEDCAC to consider what is the shortest period of impaired quality of life, well-being, or life impact that would be meaningful to detect and address, were it able to be measured accurately, and if there were an intervention available to avoid it.

Despite the progress in oncology therapies during the last several decades, there remains significant unmet medical needs for many patients. Most of the new therapies being developed are intended improve the poor prognosis in these patients but may have unintended side effects that could be measured in weeks, days, or even hours. Conversely, these new therapies could also alleviate debilitating side effects from older treatments. The resulting clinical decision making process is therefore complex and must weigh some severe events that have a duration of days or weeks against the benefits that could take months or even years to manifest.

Given the variable duration of events that may impact PROs and treatment decision making, it is highly unlikely that the selection of a PRO instrument or the adequate duration of an impaired or enhanced patient state will ever be limited by trial length in any trial having sufficient length to be useful in the oncology setting. We thus believe that Question 4 is unnecessary.

Question 5

How confident are you that PRO assessments can provide meaningful results when studied with each of the following control populations?

- *patient him/herself, before and after intervention*
- *usual care versus protocol-driven intervention*
- *historical control*

This is a very interesting question that would provide helpful guidance for the future designers of trials and supplementary data collection studies.

It is important to note that the controls listed in the questions are not mutually exclusive. For example, a single-arm oncology study could use the patient as their own control, while also comparing the magnitude, velocity, or duration of PROs from historical controls. Similarly, in controlled studies, one can look at both the group mean effect from two or more interventions, as well as individual patient responses.

One particularly interesting topic for discussion may be regarding the increased use of single-arm studies and conditional approvals. Depending on the specific patient populations and the unmet medical needs, evidence may increasingly be collected from single-arm study designs that are open-label, or other adaptive designs. It will ultimately be a regulatory review question whether there is sufficient reason to believe that the PRO data collected from such settings are appropriate for regulatory consideration or labeling. Ultimately, because the potential for bias will be of concern to regulatory agencies, and the amount of data comparing controlled and blinded PRO data to single-arm data will be limited, the regulatory bar will be high. Nevertheless, the more these studies are performed, the greater the confidence there will be to interpret the results.

As noted above, use of PRO data for additional decision making in multiple treatment settings will almost always exceed regulatory capacity and willingness to recognize such effects in labeling, or the instrument developer/user's ability to anticipate, analyze and publish on hundreds of potential scenarios. PRO effects recognized in labeling and review should be afforded substantial credence given the rigor of this setting, and thus will likely be more widespread and generalizable than labeling suggests because of the rigor by which they were originally captured and scrutinized.

Conclusion

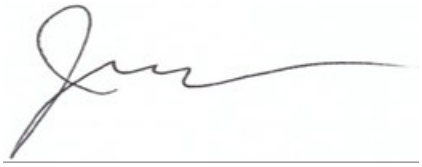
Amgen supports CMS's convening of a meeting of the MEDCAC to assess and provide recommendations on the incorporation of existing PRO assessment tools into future clinical studies. We agree that more conversations need to be had regarding the assessment of the patient-specific effects of medical treatments. This is a rapidly evolving area of research, and there is a risk that use of current regulatory standards for PRO validity and labeling will suppress clinically appropriate extrapolation of effects beyond the confines of the label and hamper further research and use of high quality PRO instrumentation in more settings. Encouraging wider use of PROs for decision making and treatment in more settings is important

because this has a sustaining and reinforcing effect on the validity of PRO results in the entire area of oncology.

Because of the potential value of PROs and the specificity of many trials, we believe it is important that the voting questions for the MEDCAC encourage a broad and nuanced approach to the value and generalizability of PROs.

We thank CMS and MEDCAC for the opportunity to comment on this important topic. Please contact me by phone at (202) 585-9659 or by email at jspangle@amgen.com if you have any questions.

Regards,



Jason Spangler, MD, MPH, FACPM