Proposed Clinical Endpoints Guidance: Knee Osteoarthritis

Document Issued on June 22, 2023

Summary of Comments Received (received July 20, 2023 – August 21, 2023)

CMS received nine comments on the proposed guidance posted on June 22, 2023. This Appendix to the final guidance summarizes and responds to the major themes of the public's comments. Commenters included four device manufacturers, a device manufacturing trade group, two academic medical centers, one physician specialty society, and an individual without affiliation. In general, commenters supported the Clinical Endpoints Guidance (CEG) goals of providing "a framework for more predictable and transparent evidence development" through the identification of the outcomes from treatment of knee osteoarthritis (OA) that would be most helpful in determining coverage policy. Some submitters expressed concerns about the process used to generate the guidance and specific aspects of the guidance.

Comments on the CEG Process

Topic Selection: Several commenters requested that CMS publish the criteria by which CEG topics will be selected and lists of upcoming topics with expected dates of proposed guidance. Commenters also requested the opportunity to have input into topic selection.

> Response: CMS hopes to issue several CEG documents each year. CMS appreciates manufacturers' and investigators' need to know early in evidence development the type of formal guidance that might be forthcoming. However, the CEG program is in its early stages, and the mechanism and criteria for topic selection are still evolving. In general, the CEG program will target clinical areas with very active research and development and those where we are aware of important uncertainty about appropriate endpoints. The FDA Breakthrough Devices Program may be one source that may be used to identify areas where a CEG could be helpful. Applicability to large segments of the Medicare population would also be an important consideration. We will provide more explicit criteria as CMS gains experience with the first several CEG documents. All CEG documents will be posted for public comment before finalization. Where the literature does not reveal a consensus on core outcomes, CMS may refer the topic to the Medicare Evidence Development & Coverage Advisory Committee (MEDCAC)[1] to clarify the most important outcomes and their clinically meaningful differences.

Stakeholder Involvement: Several commenters urged broad stakeholder engagement in all stages of CEG development. They also stressed the importance of engaging with patients, their advocates, and device manufacturers early on. Some commenters envisioned negotiations between research sponsors and CMS and requested that a "least burdensome" approach be maintained.

> Response: CMS expects to publish more detail on the CEG process once we have gained more experience making these guidance documents most valuable and efficient. All CEG documents will include a public comment period, and some will also be referred to the MEDCAC to solicit

[1] For more information, see: https://www.cms.gov/Regulations-and-Guidance/Guidance/FACA/MEDCAC

stakeholder feedback. When developing CEGs, CMS aims for a streamlined process in which multiple CEGs may be completed each year so that a library of guidance accumulates relatively quickly. Using published clinical research and consensus statements to identify core outcomes, we can ensure that proposed outcomes are consistent with clinical and investigative feasibility. Although guidance documents will be specific to therapeutic areas, cross-cutting characteristics of the kind of outcomes that are most useful to Medicare coverage decisions will emerge and may be extrapolated to topic areas that have not yet been reviewed. CEG documents will not dictate a required list of outcomes that must be studied to satisfy the reasonable and necessary standard. Nor do CEGs prescribe specific measurement instruments. Instead, CEGs are intended to clarify the most relevant measures and instruments for demonstrating improved health outcomes and satisfying CMS coverage requirements.

Clinical Evidence Review: A few commenters described what they saw as weaknesses in the methodology used to prioritize outcomes in the Clinical Endpoints Review (CER), which is a systematic review of the published literature evaluating outcomes or reports of consensus statements based on an established process. They suggested that citation frequency should not be the only source of recommended domains and outcomes (see *Recommended Endpoints and Measurement Instruments* under **Comments on CER Findings and CMS Conclusions**). The specialty society argued that "several of the included studies included heterogeneous patient populations, such as patients with rheumatoid arthritis (RA) or other musculoskeletal conditions in addition to OA . . . it would be best to include only those studies that include the target population."

Response: CMS appreciates these thoughtful observations. CMS has again reviewed the included studies to ensure their inclusion in the CEG remains appropriate. However, since the purpose of the CER was to identify outcomes that are commonly measured rather than to assess the effectiveness of specific technologies, heterogeneous patient populations do not pose a problem as long as the studies were evaluating technologies that are appropriate for treating OA and recruited a substantial proportion of patients with OA. In the study that included RA patients, they comprised only 1.5% of the study group; the remaining study participants had OA in the study that included a mix of musculoskeletal conditions involving the knee (Hofstede et al., 2015). From a systematic review of studies evaluating outcome measures for various knee conditions, the CER included only instruments used in patients with knee or knee/hip OA.

Comments on CER Findings and CMS Conclusions

Generalizability to the Medicare Population: One device manufacturer asked that CMS be more explicit about the percentage of Medicare beneficiaries that should be included in a trial and the subpopulations that should be adequately represented. The comment suggests that "the percentage of Medicare beneficiaries enrolled in clinical trials . . . be roughly proportional to their representation in the disease state under investigation and at least adequate (i.e., large enough) to provide reasonable assurance of safety and effectiveness in subgroup analysis using interaction testing".

Response: In general, in order for an item or service to be covered under Medicare, it must meet the standard described in section 1862(a)(1)(A) of the Social Security Act – that is, it must be reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member. When making coverage determinations, CMS policies have long considered whether the item or service is not just safe and effective but also whether

the item or service is not experimental or investigational and is appropriate for Medicare beneficiaries.[2] When making this determination, CMS generally requires that evidence from clinical studies apply to the intended recipients of the Medicare population(s). Applicability assessment depends on whether a new technology's effectiveness would reasonably be expected to vary between the populations studied in clinical trials and the most likely Medicare recipients, who are often older and have more comorbidities.

Data Sources and Study Designs: One academic medical center and one device manufacturer requested guidance on data sources and data collection strategies for the recommended outcomes, particularly concerning the generation of real-world evidence and the design of fit-for-purpose studies.

Response: CMS appreciates these concerns about real-world evidence and fit-for-purpose studies and recognizes that some forms of real-world evidence, such as administrative claims and electronic medical records, will often need to be supplemented with data that yields patient reported outcomes (PROs) and additional clinical measurements. For example, some outcomes may be collected through registry-based studies and pragmatic randomized controlled trials. A discussion of potential data sources and data collection strategies for real-world data is beyond the scope of this CEG. However, CMS expects to publish proposed fit-for-purpose study guidance in the coming months.

Recommended Endpoints and Measurement Instruments: A few commentators requested more specific guidance on endpoints and instruments. For example, the literature reviewed in this CEG did not identify specific instruments for every domain. The specialty society recommended a range of motion to measure the stiffness domain and advised that 'Radiographic imaging' as an instrument for the joint structure domain was too vague and should be replaced with a specific imaging target, such as evidence of disease progression.

One device manufacturer asked CMS to provide more specific guidance regarding endpoints that CMS considers validated and well-established, along with corresponding minimal clinically important differences (MCIDs). They also requested that we provide guidance for manufacturers that intend to use endpoints not included in the CEG. For example, the commenter cited the FDA's endorsement of specific tools in its Qualification of Medical Device Development Tools guidance document. One academic medical center pointed out that neither the visual analogue scale (VAS) nor a Likert scale are validated instruments for assessing patient satisfaction in knee OA.

Some commenters urged CMS to include additional outcomes or endpoints in the prioritized list of domains beyond those identified in the articles that met the inclusion criteria of the CER. One device manufacturer noted that the Forgotten Joint Score, the Knee Society Score, and the Oxford Knee Score are commonly used measures that should be included in the guidance. One specialty society comment suggested that cost and accessibility of treatment be added as domains under the Resource Use and Economical Impact section. Two academic medical centers urged the inclusion of patient satisfaction as a standard measure in CEG documents, whether the contractor's methodology led to its inclusion or not.

[2] Medicare Administrative Contractors (MACs) determine if evidence exist to consider and item or service to be reasonable and necessary if the MAC determines that the service is safe and effective, not experimental or investigational, and appropriate. For more information see the CMS Program Integrity Manual, Chapter 13.5.4, available at https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/downloads/pim83c13.pdf..

Response: To help clarify that information from radiographic imaging would be considered outcome measures, the text in the final document has been modified to include a description of the Kellgren and Lawrence radiographic grading system.

In the interest of producing flexible guidance that may also apply to new measurement instruments that may emerge, CMS has not recommended specific endpoints and instruments within each domain. However, CMS understands the commenters' desire for concrete examples. The CEG for Knee Osteoarthritis provides numerous examples that could be used within some of the prioritized domains. Still, these lists are not intended to be exhaustive, and CMS recognizes that the CER may have missed publications that support additional endpoints. As stated in the CMS Conclusions section of the proposed and final guidance, "CMS does not recommend any specific clinical endpoint or instrument that was identified in this review, but generally recommends that clinical studies include a range of outcomes that reflect multiple attributes of an item or service within a clinical study" and "When choosing among the available clinical endpoint options, CMS recommends that clinical studies prioritize validated endpoints/instruments and those with well-established/published minimal clinically important differences (MCID) values." The latter statement assumes that the use of endpoints not explicitly named in CEG documents should be based on published validation studies or consensus statements derived from systematic methods such as Delphi panels where possible. Nonetheless, CMS recognizes that not all widely used instruments within a given disease area have been formally validated.

CMS has considered the three instruments that appeared to be missing from the guidance document. The Forgotten Joint Score did not appear in any of the publications that met the inclusion criteria for the CER, which stipulated that the publication had either a) implemented a systematic search of the literature to evaluate outcomes or b) used an established process (e.g., Delphi) to arrive at consensus on outcomes. Some of the selected publications that cited the Oxford Knee Scale (OKS) or Knee Society Score (KSS) did not specify whether the entire scale or one of the subscales was being considered; those publications were thus originally not counted toward overall citation volume. In the final guidance document, all publications citing one of these two instruments, regardless of how the instrument was used, have been counted toward citation volume. The OKS and KSS now appear in Table 2 (Prioritized Outcome Domains and Instruments).

CEG documents are intended to clarify the type of evidence needed when making coverage determinations to demonstrate that an item or service is reasonable and necessary. CMS does not consider cost when reviewing the effectiveness of technologies in the Medicare population and thus does not recommend cost-related measurements to establish whether an item or service is reasonable and necessary. Nevertheless, CMS is committed to ensuring equitable access to covered services.

The CEG series intends to establish a standard and objective process for identifying commonly used clinical endpoints and their clinically meaningful differences. Thus, it will avoid making a priori conclusions about specific outcome measures or domains. Nonetheless, CMS coverage decisions generally prioritize health outcomes that are important to patients, that is, outcomes

that reflect what patients experience as the result of treatment. We note that patient satisfaction is included as a domain in Table 2 of the CER.

Identification of Domains: One of the academic medical centers and one specialty society expressed confusion over the difference between the two domains listed in the CER under 'Life Impact': 'Physical function' and 'Function/functional ability.' The specialty society questioned why outcomes such as return to work or return to sports were envisioned as 'Function/functional ability.'

Response: CMS has made clarifying changes in the final CEG for knee osteoarthritis. These changes include a substitution of "Role function" for "Function/functional ability" and the addition of an explanation distinguishing between role function (ability to function in society) and physical function (ability to perform physical activities such as walking or climbing stairs).

Follow-up Intervals: Device manufacturers submitted most comments about follow-up interval recommendations. Some commenters emphasized that clinically meaningful follow-up intervals vary by health condition and treatment types and thus should be determined on a case-by-case basis rather than defined in a guidance document. One commenter argued that in some cases, the value of a new technology is that it accelerates the achievement of important health outcomes such as a return to sports or work without necessarily altering long-term outcomes. They argued that studies with short follow-up intervals may be appropriate in those cases. Other commenters argued that follow-up intervals should depend on trial study design, study power, and statistical plan or requested that CMS supply clear stopping rules for data collection. One specialty society suggested a minimum of two years of follow-up for evaluating the durability of implanted devices such as joint replacements but pointed out that ten or more years may be necessary to assess durability for younger patients undergoing joint replacement.

Response: CMS agrees that a guidance document cannot address all the factors that could apply to all new technologies. In the example of a new technology that accelerates benefits, CMS would still need evidence that long-term benefits are comparable to those of other covered treatments but would consider evidence of accelerated benefit. The appropriateness of follow-up intervals relates to the durability of effectiveness and the potential for low-frequency adverse effects to occur over longer time frames; therefore, CMS believes that meaningful follow-up intervals should dictate study design, sample size, and statistical plan considerations, rather than the other way around. CMS appreciates the clinical perspective behind the suggestion that extended follow-up may be needed to assess durability in some populations. However, while long-term follow-up may be valuable to the evidence base, CMS must often make coverage determinations well before such long-term studies can be completed.

Alignment with Other CMS Programs: One academic medical center urged consistency between the measurement instruments prescribed in CMS value-based payment programs, such as the Merit-Based Incentive Payment System (MIPS) and the Comprehensive Care for Joint Replacement (CJR) Model. The commenter stated that the Knee injury and Osteoarthritis Outcome Score (KOOS) is the only outcome listed in the CEG for Knee OA that is regularly collected in other CMS programs.

Response: CMS does not expect any one study of a knee OA treatment to address every outcome in the prioritized list generated in a CEG. The CEG for Knee Osteoarthritis recommends only that "investigators consider the 11 prioritized outcome *domains* identified in Table 2 of this technology assessment when designing knee osteoarthritis clinical studies" and states that "CMS

does not recommend any specific clinical endpoint or instrument that was identified in this review, but generally recommends that clinical studies include a range of outcomes that reflect multiple attributes of an item or service within a clinical study." Metric-based programs such as MIPS and the CJR Model must assure comparability across providers by requiring prescribed measurement instruments. In the context of clinical studies designed to support coverage decisions, measurements of multiple outcomes across several domains are needed to provide a balanced picture of the value of a new technology.

Study Designs: An academic medical center and a device manufacturer requested more detailed guidance regarding study designs while urging flexibility to accommodate novel or innovative designs.

Response: Beyond generally recommended follow-up intervals for different outcome types, study design issues are beyond the scope of the CEG series. Further information on study design considerations is detailed in the Coverage with Evidence Development and the CMS National Coverage Analysis Evidence Review Guidance documents.

Other: One commenter with no stated affiliation expressed concerns about guidelines for the use of bracing.

Response: In keeping with the intent of the CEG program, the CEG for Knee OA does not address specific technologies or provide any guidance on treatment appropriateness.