PLUGS Consensus Recommendations Regarding the Development of Medical Necessity Policies, Version 1.0

3/17/2025

Authors: PLUGS Insurance Alignment Committee, Medical Necessity Working Group

Corresponding Author: Michael Astion, MD, PhD: michael.astion@seattlechildrens.org

Background:

A medically necessary laboratory test refers to a test needed to diagnose or manage a health condition (National Association of Insurance Commissioners (NAIC), 2024). Medical necessity determinations should support commonly accepted standards of care in the community (NAIC, 2024) and should not support excessive testing beyond what is medically reasonable (Center for Medicare and Medicaid Services (CMS), 2024a). Medical necessity policies state the clinical criteria for a test to be medically necessary, and they include the evidence supporting the criteria. Payers prefer basing medical necessity policies on peer-reviewed evidence graded on strength, as well as guidelines published by the government or mainstream clinical societies (Astion, 2023). In practice, for laboratory testing, payers often must rely on guidelines based on weaker evidence, including expert consensus. Besides guidelines and peer-reviewed evidence, payers use claims data, opinions from their own board-accredited specialists, and feedback from providers to inform medical necessity policies.

One goal of medical necessity policies is to block fraud, waste, and abuse (FWA). CMS has defined these terms (CMS, 2016; CMS, 2021). In the context of medical necessity policies for lab tests, abuse refers to billing for medically unnecessary tests, usually without the intent to deceive to gain payment. Waste refers to misutilization, especially test overutilization. Fraud is waste or abuse with the intent to deceive to gain payment (CMS 2016; CMS, 2021).

There are no universally accepted standards for medical necessity policies (NAIC, 2023; Astion, 2023). Different payers produce documents of varying size, contents, structure, educational level, and accessibility. Unfortunately, it is common that inclusion and exclusion criteria vary for the identical test. This is frustrating, inefficient, and costly to care providers and clinical laboratories, who are asked to manage multiple, complex payer policies and procedures, within the same clinical practice.

The current trends are for payers to manage laboratory test utilization through a focus on medical necessity criteria, and to increase annually the number of tests

under management. This test management comes either directly from payers, who employ experts to develop medical necessity policies, or from laboratory benefits managers (LBMs) that service the payers (Phillips and Deverka, 2019).

The two most common methods to manage laboratory testing through determining necessity are pre-authorization and post-service claims processing algorithms. Pre-authorization produces a medical necessity determination or coverage decision before the test is provided, potentially offering patients and labs financial protection (CMS, 2024b). Post-service claims processing is riskier for patients since patients often must pay if the test is denied, and if they are not required to pay or unable to pay, the lab bears the cost of performing an unreimbursed test. Usually, pre-authorization is applied to more expensive tests, for example genomes or exomes ordered for the evaluation of inherited diseases. Automated claims processing algorithms tend to focus on high volume tests including the most common tests, like vitamin D, thyroid testing, lipid panels, and respiratory virus panels. Typically, these tests are inexpensive. The claims processing algorithms determine medical necessity by matching lab CPT codes, patient demographic information, and allowable or deniable ICD-10 codes.

The growth and competition in laboratory benefits management is associated with increased denials caused by stricter medical necessity criteria, which are based on the requirement for higher levels of evidence. Ideally, evidence-based medicine alone would be the foundation for medical necessity decisions. Unfortunately, the current medical system is not ideal, and an over-reliance on exacting standards of evidence produces two significant problems for patients, providers, and labs (Astion, 2023). First, payers and LBMs, as well as professional societies and governmental agencies, vary in their evidence interpretation leading to significant variation in coverage for identical tests (Bauchner and Ioannidis, 2024). When evidence is weaker, as often is the case for less common diseases, payers may not allow testing even though it is supported by one or more guidelines.

The second problem is that the standard of care and evidence-based medicine overlap, but are not identical (Astion, 2023). The standard of care is used as both a medical and legal term and has a range of definitions (Moffett and Moore, 2011). A composite definition of the standard of care is the expectation of the average provider to diagnose, treat, monitor, and communicate about a health condition. Standards of care in laboratory testing are often based on weaker evidence from small case control studies, observational studies, or a consensus of academically oriented, board-certified medical specialists. Larger well-controlled studies and randomized control trials are less common and tend to be restricted to the highest volume tests for common diseases. In practice, the legal standard of care comes from experts, and their opinion is based on peer-reviewed research; guidelines, practice updates, and other educational documents from professional societies and the government; textbooks and online information from medical publishers; and historical practice patterns. Significant deviation from the standard of care is the basis of malpractice lawsuits. Patients and providers believe insurance policies should support the standard of care and are frustrated if denied a test, which was

ordered to meet the standard of care, but which was denied by the payer due to their interpretation of the evidence (Astion, 2023).

The purpose of the recommendations presented here are to enable the development of medical necessity policies that reduce FWA while providing flexibility in medical practice and decreased administrative burden. It is our hope that these recommendations are adapted by payers and LBMs and modified over time through feedback. Eventually, we hope they become voluntary standards.

Recommendations:

1. **Create medical necessity policies that allow a broad path of reasonable care, within which providers may practice care that is safe and unlikely to cause FWA, and outside of which represents significant risk of harm or FWA.**

   Our PLUGS consensus committee refers to this approach as "Guardrails" (Astion, 2023), and it is illustrated in Figure 1. Guardrails block a significant amount of FWA while leaving sufficient room for providers and labs to practice, especially in support of the most severely ill patients. Since its inception, PLUGS has applied guardrails to policy development and review. Overutilization encompasses about 20% of lab testing (Zhi et al., 2013; Kroner, et al., 2022). Guardrails reduce overutilization, while supporting denials for all CPT codes representing obsolete tests, duplicate tests, or tests with no evidence of clinical utility.

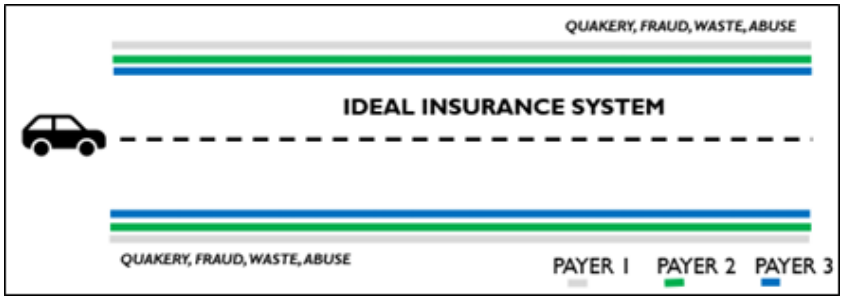

Figure 1. The guardrails approach to medical necessity policies. In this example, the payers have policies that block significant cases of quackery, fraud, waste, and abuse. Payer 3 is the strictest but nonetheless allows sufficient flexibility to practice medicine in a multi-payer environment. Quackery refers to promotion of a useless medical service with profitable intent (Quackwatch, 2001).

Guardrails encompass a tradeoff between two contrasting truths. The first is that without interventions, government and private payers will pay for a significant amount of FWA. This is proven by an analysis of claims data, as well as data available through court cases brought by commercial payers or the United States Department of Justice (USDOJ) (USDOJ, 2022; USDOJ, 2015; Tycko and Zavareei LLP, 2024). The second truth is that most providers are competent and typically practice within guardrails. Physicians tend to practice outside guardrails for difficult cases such as rare diseases; highly morbid cases intractable to treatment; and complex patients with multiple, complex comorbidities. These cases benefit

from flexibility in the medical necessity policy to allow care in cases where there is clinical reasoning that is specific and logical.

Guardrails are consistent with the approach used in clinical guidelines that inform medical necessity policies and the standard of care. Clinical guidelines tend to outline a spectrum of reasonable practice. Specific guidance is given when the evidence is strong and is an excellent fit for a clinical case. When evidence is weaker or the clinical cases are more complex, flexible language such as "may consider," "may be appropriate," or "is recommended," is preferred and purposefully utilized to allow providers more choices.

The guardrails approach acknowledges that it is abrasive for providers and labs to adjust to high variation in policies from their multiple payers (Figure 2). Variation between payers is frustrating, inefficient, and expensive for patients, labs, and providers. In practice, the variation causes hardship because patients avoid testing due to costs (Goozner, 2019; National Cancer Institute, 2024). Thus, access to testing is often determined by the patient's insurance.

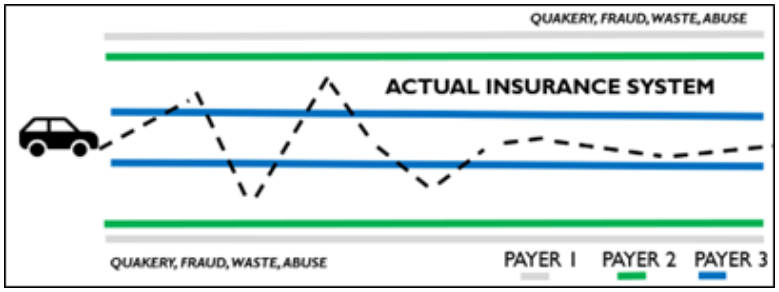

Figure 2. Significant variation in payer policies leads to different frequency of claims denials, which produces inefficiency, expense, and frustration for patients, providers, and clinical labs. In this example, payer 1 is providing guardrails and payer 3 is forcing practice down too narrow a path, with insufficient allowance for variation in clinical needs.

2. **Use wider guardrails for patients with multiple, clinically severe diagnoses and comorbidities with an emphasis on allowing higher frequency use of common tests.**

   Typical scenarios requiring wider guardrails involve extremely ill patients with multiple diagnoses and comorbidities. These patients often move between specialists, sometimes in different health systems, and there can be long intervals between evaluations by a particular specialist. The diagnoses may be common like diabetes, atherosclerotic heart disease or cancer. The most challenging situations involve rare diagnoses such as inherited diseases with severe phenotypes, or multi-symptom syndromes as can occur in autoimmune diseases, inflammatory diseases, diseases related to chronic environmental or workplace exposures, and in patients who receive organ transplants. For rare inherited diseases, a broader definition of clinical utility, which includes the value of ending the diagnostic odyssey, has been proposed (ACMG, 2015).

The first reason for broader guardrails is evidence weakens as patients become more complex. This is because the likelihood of controlling for confounding conditions decreases as patient complexity increases. The decreased ability to crosswalk results from the scientific literature to a complex patient, argues for giving providers flexibility to re-evaluate and monitor patients more frequently with a larger group of conventional tests.

The second reason for broader guardrails is the inaccuracy in coding complex patients and encounters (Horsky et al., 2018; Wei et al., 2020; Nashed et al., 2021; Schaefer et al., 2022). For these patients the combination of CPT codes, ICD 10 codes, and demographic information falls short in describing the patient encounter relative to the gold standard, which is a chart review that includes medical, surgical, and nursing notes, pharmacy data, and reports from lab, radiology, nuclear medicine, and other diagnostic procedures. In addition, a longitudinal set of claims provides a more accurate assessment of the patient than a single claim. The financial burden of coding deficiencies can fall disproportionately on patients, clinical labs, or the payer, depending on the case.

The recommendation for wider guardrails still supports the denial of tests which are fraudulent or for which there is no evidence of clinical utility. Such tests do not benefit any patients under any conditions.

3. **Write policies using simple, straightforward language.**
Ideally, policies would use simple language that could be understood by individuals without a medical background. Thus, medical jargon and complex terms would be minimized, and policies would be readily translated into multiple languages. A more realistic alternative is to write a simple summary of the policy suitable for individuals without a medical background. This is easy using artificial intelligence aids. The summary could include a glossary or parenthetical explanations for technical terms. Concrete examples could be included to illustrate when a test is considered medically necessary so that patients comprehend how the policy applies to them. Common scenarios where the test is denied could also be included with an explanation to help patients understand the reasoning behind these denials.

4. **Make it is easy for the patient, their provider, and the clinical lab to locate the medical necessity policy for a specific test. For larger policies covering multiple tests, make it easy to find the test within the policy.**

   Finding a policy is easier if the payer website is user friendly, and access to the policy can be gained without requiring a username and password. Toward that end, several payers make their policies freely available.

   Frequently, multiple tests are combined within one payer policy. For example, this is often the case with genetic tests for inherited diseases, cancer genomic profiling, tumor marker tests, and tests for cardiovascular disease risk. It can be impossible for patients or providers to find the test they are looking for within a policy and determine relevant medical necessity criteria. Artificial intelligence applications could make it easier to find each test within a policy, and the coverage criteria that applies to each test.

5. **Update medical necessity policies at least annually and highlight the changes in the updated policy. After the policy update, migrate the policy into the claims processing system with minimal delay.**

   Annual review of new evidence or changes to the standard of care offer a chance for a change in coverage. The clinical criteria in medical necessity policies must be programmed into the payer's claims processing system, and ideally this would be accomplished within one month of completing the revised policy. However, this programming tends to occur with a 3 to 12-month delay following policy changes. In cases where medical necessity criteria have broadened, and a previously denied test is going to be allowed, patients will be denied during the delay and responsible for payment. If the patient cannot pay, the lab will bear the cost of providing the test without compensation.

6. **Make claims processing logic transparent to laboratories, patients, and providers.**

   In an ideal health system, ethical providers and labs deserve access to the proper ICD-10 coding and demographic information required for tests they deem medically necessary. It would be helpful to know the ICD-10 codes that are commonly covered or commonly denied. Unfortunately, the current commercial payment system is thwarted by two problems hindering this recommendation. First, unscrupulous providers and labs can use knowledge of the claims processing logic to gain approval for claims previously denied as FWA. Second, claims processing systems require significant investment to develop and maintain, and are protected intellectual property. Publishing of claims logic allows competing payers to have a valuable tool with minimal investment.

7. **Establish a clear, straightforward mechanism for providers and patients to share feedback on medical necessity policies.**

   This can be an online form or email address used to collect feedback that will inform policy improvements. This will provide a more abundant and nuanced set of opinions than can be gathered from appeals alone. In addition, including patients and providers in the policy process values their experiences and concerns, enabling a better alignment between payers, patients, providers, and labs.

8. **Only create a medical necessity policy if it is demonstrably necessary to block FWA and does not exacerbate underutilization of a test.**

   The payer should provide evidence from claims data or the peer-reviewed literature that the laboratory test has resulted in FWA. In addition, payers should determine if the administration of a policy to reduce FWA exacerbates underutilization of tests. There are variety of tests that are both underutilized and overutilized. Testing for celiac disease, and the monitoring of diabetes by HbA1C are two examples. Development and administration of the medical necessity policy should only proceed if underutilization is reduced or unaffected.

9. **Create medical necessity policies that are not contradictory to other policies, including pharmacy pathways or policies.**

   Patients, providers, and labs find it frustrating and unfair if a test, for example a cancer genomic profile of a tumor, is deemed necessary for guiding treatment in a pharmacy policy, but experimental, investigational, and unproven in a lab policy.

10. **The medical necessity policy states the evidence review methods.**

    The intention of a medical policy is to determine if there are clinical scenarios for which a laboratory test may be useful, and if so what level of evidence supports that use. One key tool in this determination is a literature search and review. For the literature search, payers should provide the keywords, search strings, and dates of the search. For the literature review, there should be proper citation of the evidence reviewed, and a description of the methods regarding how that evidence was identified and graded.

    It is not required or expected that the evidence threshold for determining medical necessity be the same for every test. For high frequency tests in the setting of common diseases, for example lipid testing for atherosclerotic cardiovascular disease, thousands of research studies and hundreds of systematic reviews would have been conducted, and therefore testing practices associated with the standard of care are likely to be supported by high grade evidence. In contrast, for rare diseases and other uncommon clinical situations with high morbidity and mortality, the evidence standard may need to be relaxed to accommodate the standard of care. This may include relying on guidelines based on consensus of board-accredited specialists.

References:

What is Medical Necessity? 2023. National Association of Insurance Commissioners (NAIC). https://content.naic.org/sites/default/files/consumer-health-insurance-what-is-medical-necessity.pdf Accessed March 17, 2025.

Items and Services Not Covered Under Medicare. 2024a. CMS-Medicare Learning Network Booklet 906765. https://www.cms.gov/outreach-and-education/medicare-learning-network-mln/mlnproducts/downloads/items-and-services-not-covered-under-medicare-booklet-icn906765.pdf Accessed March 17, 2025.

Astion M. 2023. Medical Necessity Policies: Creating Opportunities for Better Test Utilization and Fair Payment. *Clinical Laboratory News*. 49(7): Available at: https://www.aacc.org/cln/articles/2023/julaug/medical-necessity-policies-creating-opportunities-for-better-test-utilization-and-fair-payment Accessed March 17, 2025.

Fraud, Waste, and Abuse Toolkit. Health Care Fraud and Program Integrity: An Overview for Providers. 2016. Center for Medicare and Medicaid Services (CMS).

https://www.cms.gov/files/document/overviewfwaprovidersbooklet072616pdf Accessed March 17, 2025.

Medicare Fraud & Abuse: Prevent, Detect, Report. 2021. Center for Medicare and Medicaid Services (CMS). https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/Downloads/Fraud-Abuse-MLN4649244.pdf Accessed March 17, 2025.

Phillips KA, Deverka PA. The Emerging Use by Commercial Payers of Third-Party Lab Benefit Managers for Genetic Testing. Health Affairs. October 23, 2019. https://www.healthaffairs.org/content/forefront/emerging-use-commercial-payers-third-party-lab-benefit-managers-genetic-testing Accessed March 17, 2025.

Prior Authorization and Pre-Claim Review Initiatives. 2024b. Center for Medicare and Medicaid Services. https://www.cms.gov/data-research/monitoring-programs/medicare-fee-service-compliance-programs/prior-authorization-and-pre-claim-review-initiatives Accessed March 17, 2025.

Bauchner H, Ioannidis JPA. The Subjective Interpretation of the Medical Evidence. JAMA Health Forum. 2024 Mar 1;5(3): e240213. doi: 10.1001/jamahealthforum.2024.0213. PMID: 38551587.

Moffett P, Moore G. The standard of care: legal history and definitions: the bad and good news. West J Emerg Med. 2011 Feb;12(1):109-12. PMID: 21691483; PMCID: PMC3088386.

Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaout R. The landscape of inappropriate laboratory testing: a 15-year meta-analysis. PLoS One. 2013;8(11): e78962. Published 2013 Nov 15. doi:10.1371/journal.pone.0078962

Kroner GM, Richman S, Fletcher A, Dickerson J, Jackson BR. Current state of laboratory test utilization practices in the clinical laboratory. Acad Pathol. 2022;9(1):100039. Published 2022 Aug 9. doi:10.1016/j.acpath.2022.100039

Quackery, Fraud, and "Alternative" Methods: Important Definitions. 2001. Quackwatch. https://quackwatch.org/related/quackdef2/ Accessed March 17, 2025.

Justice Department Charges Dozens for $1.2 Billion in Healthcare Fraud. July 20, 2022. US Department of Justice. https://www.justice.gov/opa/pr/justice-department-charges-dozens-12-billion-health-care-fraud Accessed April 15, 2024

Millennium Health Agrees to Pay $256 Million to Resolve Allegations of Unnecessary Drug and Genetic Testing and Illegal Remuneration to Physicians. 2015. US Department of Justice. https://www.justice.gov/opa/pr/millennium-health-agrees-pay-256-million-resolve-allegations-unnecessary-drug-and-genetic Accessed May 17, 2025.

Tyco and Zavareei Whistleblower Practice Group. 2024. Not Positive: Last Week's Lab Fraud Schemes. The National Law Review. https://natlawreview.com/article/not-positive-last-week-s-lab-fraud-schemes#google_vignette Accessed May 17, 2025.

Goozner M. Editorial: An epidemic of financial toxicity. Modern Healthcare. August 30, 2019. https://www.modernhealthcare.com/opinion-editorial/editorial-epidemic-financial-toxicity Accessed March 17, 2025.

Financial Toxicity (Financial Distress) and Cancer Treatment. 2024. National Cancer Institute. https://www.cancer.gov/about-cancer/managing-care/track-care-costs/financial-toxicity-pdq Accessed March 17, 2025.

ACMG Board of Directors. Clinical utility of genetic and genomic services: a position statement of the American College of Medical Genetics and Genomics. Genet Med. 2015 Jun;17(6):505-7. doi: 10.1038/gim.2015.41. Epub 2015 Mar 12. PMID: 25764213.

Horsky J, Drucker EA, Ramelson HZ. Accuracy and Completeness of Clinical Coding Using ICD-10 for Ambulatory Visits. AMIA Annu Symp Proc. 2018 Apr 16; 2017:912-920. PMID: 29854158; PMCID: PMC5977598.

Wei MY, Luster JE, Chan CL, Min L. Comprehensive review of ICD-9 code accuracies to measure multimorbidity in administrative data. BMC Health Serv Res. 2020 Jun 1;20(1):489. doi: 10.1186/s12913-020-05207-4. PMID: 32487087; PMCID: PMC7268621.

Nashed A, Zhang S, Chiang CW, Zitu M, Otterson GA, Presley CJ, Kendra K, Patel SH, Johns A, Li M, Grogan M, Lopez G, Owen DH, Li L. Comparative assessment of manual chart review and ICD claims data in evaluating immunotherapy-related adverse events. Cancer Immunol Immunother. 2021 Oct;70(10):2761-2769. doi: 10.1007/s00262-021-02880-0. Epub 2021 Feb 24. PMID: 33625533.

Schaefer JW, Riley JM, Li M, Cheney-Peters DR, Venkataraman CM, Li CJ, Smaltz CM, Bradley CG, Lee CY, Fitzpatrick DM, Ney DB, Zaret DS, Chalikonda DM, Mairose JD, Chauhan K, Szot MV, Jones RB, Bashir-Hamidu R, Mitsuhashi S, Kubey AA. Comparing reliability of ICD-10-based COVID-19 comorbidity data to manual chart review, a retrospective cross-sectional study. J Med Virol. 2022 Apr;94(4):1550-1557. doi: 10.1002/jmv.27492. Epub 2021 Dec 8. PMID: 34850420; PMCID: PMC9015484.