PUBLIC COMMENT

Patient Safety Requires Meaningful Human Judgment in Software-Assisted Medical Decisions

FDA Docket No. FDA-2018-N-1910

Development of 21st Century Cures Act Section 3060 Required Report: Request for Input

Submitted by P.A.R.T. USA
Pain Awareness - Right to Treatment
Ingrid "Kitty" Trimm, Founder / Executive Director
August 10, 2026

Executive Statement

P.A.R.T. USA submits these comments in response to FDA's request for input on the risks and benefits of non-device software functions and best practices to promote patient safety, education, and competency. Our concern is not with technology itself. Healthcare software can calculate, organize, identify patterns, reduce administrative burdens, and provide valuable decision support. The patient-safety problem begins when a tool becomes an authority.

> **When software materially affects an individual patient's diagnosis, treatment, medication access, dispensing, tapering, referral, coverage, or risk classification, the software should remain advisory rather than dispositive. A qualified human professional must retain a meaningful opportunity to independently review the software's basis, consider information the software does not contain, correct erroneous inputs, obtain additional context, override the output when appropriate, and exercise individualized clinical judgment before a consequential action is taken.** [1][2]

This position is consistent with the framework Congress and FDA already use for qualifying non-device clinical decision support. Section 520(o)(1)(E) of the Federal Food, Drug, and Cosmetic Act requires that the healthcare professional be able to independently review the basis of a recommendation so the professional is not intended to rely primarily on the software for an individual diagnosis or treatment decision. Section 520(o)(3) further directs FDA to consider the likelihood and severity of harm and whether the professional has a reasonable opportunity to review the recommendation's basis. [1]

1. A computer can process the data correctly and still be wrong about the patient

One of the most important distinctions in software-assisted medicine is the difference between computational correctness and clinical correctness. A computer can perform its programmed calculation perfectly while producing a clinically misleading result because the underlying information is incomplete, inaccurate, outdated, miscoded, or missing important context. [2]

Patient outcomes may be influenced by age, diagnosis, disease severity, organ function, body size, metabolism, genetics, medication tolerance, previous exposure, coexisting conditions, concurrent medications, treatment history, previous treatment failures, financial barriers, insurance restrictions, functional response, and many other factors. Some of those variables may exist in structured electronic data. Others may not.

FDA's January 2026 Clinical Decision Support Software guidance recognizes this problem by recommending that developers identify required patient inputs, explain their relevance, address data-quality requirements, disclose performance limitations, and make relevant knowns and unknowns visible, including missing, corrupted, or unexpected information. [2]

The accuracy of a calculation cannot compensate for the incompleteness of the picture being calculated. A sophisticated algorithm operating on incomplete information may simply produce a more sophisticated error.

2. Software should identify questions - not terminate them

Healthcare software is most valuable when it helps the clinician recognize something that requires additional investigation. A risk score might appropriately communicate: "There is information here that deserves further review." That is fundamentally different from treating the score as meaning: "This number determines what must happen to this patient."

Medicine requires questioning: Why is this medication present? Why was this treatment chosen? What alternatives were tried? What failed? What worked? Is the diagnosis accurate? Is a prescription being used for an uncommon but legitimate purpose? Is relevant information missing from the computerized record?

> **An algorithm should identify questions that need to be asked, not become a reason to stop asking questions.**

3. Medication identity does not necessarily reveal clinical context

Buprenorphine provides a clear example of why medication data must be interpreted carefully. VA materials identify transdermal and buccal buprenorphine formulations as chronic-pain treatments, while sublingual buprenorphine and buprenorphine/naloxone are principally used for opioid use disorder and may also sometimes be used off-label for pain.[10]

The presence of a medication in a prescription or dispensing record does not, standing alone, establish the full clinical reason why that medication was prescribed. Some systems may have linked diagnoses, chart information, claims information, or narrative documentation; others may not.[10]

> **Medication identity alone should not automatically be converted into a diagnosis, motive, behavior classification, or treatment conclusion without review of the relevant clinical context.**

4. MME must remain a tool - never an absolute treatment rule

Morphine milligram equivalents (MME) are a standardized calculation tool intended to help compare opioid dosages across different medications. MME should be used only as one piece of information within a broader individualized clinical evaluation. It should never become an absolute threshold for treatment.[6]

CDC's 2022 Clinical Practice Guideline states that equianalgesic conversions are estimates and cannot account for all individual variability in genetics and pharmacokinetics. CDC also warns against mechanically using calculated MME when changing opioids because incomplete cross-tolerance, opioid-specific characteristics, and individual pharmacokinetic differences matter.[6]

CDC further states that its dosage recommendations are not intended to be rigid standards or absolute limits. CDC specifically instructs healthcare administrators: "Do not set rigid dosage thresholds."[7]

A calculation can provide useful information. It cannot determine, by itself, what treatment is appropriate for an individual human being. MME should remain a clinical tool - never an absolute rule governing whether a patient receives, continues, loses, or is forced to change medically appropriate treatment.

5. PDMPs and risk-scoring systems have demonstrated potential for serious harm

P.A.R.T. USA is deeply concerned about the harms that can occur when Prescription Drug Monitoring Programs, proprietary risk scores, and related computerized systems influence medical and pharmacy decisions. These concerns are not theoretical.[11]

A 2021 mixed-methods systematic review found that PDMP information influenced clinicians' decisions, including controlled-substance prescribing, refusal to prescribe or treat, communication, referrals, care coordination, risk mitigation, and stigma, with both intended and unintended patient outcomes.[11]

When computerized information contributes to medication refusal, altered treatment, stigma, or disruption of the clinician-patient relationship, that is harmful and it is a patient-safety concern. P.A.R.T. USA is also concerned that numerical risk scores and highly visible alerts can create fear and institutional pressure that discourage doctors and pharmacists from exercising independent professional judgment.[11][12]

A score can look objective. That does not make the conclusion absolute. Computerized risk information should be treated as a reason to investigate further, not as an answer. Scores should be considered general information and an evaluation of available data. They should trigger examination, questioning, verification, and individualized assessment. They should never be treated as absolutes.

6. NarxCare illustrates why prediction is not the same as clinical authority

Some NarxCare-related scores have undergone technical or associational validation. A 2021 peer-reviewed study reported fair concurrent validity between a NarxCare Narcotic Score and the WHO ASSIST screening instrument. That evidence should be acknowledged. [14]

However, technical association is different from demonstrating that use of a score safely determines what treatment an individual patient should receive. A 2023 analysis in the Journal of the American Medical Informatics Association concluded that publicly available NarxCare Overdose Risk Score evaluations were limited and that the ORS had not undergone clinical evaluation establishing its effects on patient or provider outcomes; the authors also identified limitations in generalizability and public subgroup analysis. [13]

A score may correlate with an outcome without proving that withholding treatment because of that score improves the patient's outcome. A tool validated to estimate risk is not automatically validated to support medication denial, pharmacy refusal, involuntary tapering, coverage denial, dismissal from care, referral refusal, or another consequential intervention. [13]

Bamboo Health states that NarxCare is intended to aid rather than replace medical decision-making and should not be the sole justification for providing or refusing medication. Its March 2026 ORS documentation also states that the score does not predict whether a particular patient will experience an unintentional overdose death now or in the future. [15][16]

7. Automation bias means simply putting a human in the workflow is not enough

Human oversight must be meaningful, not ceremonial. FDA's 2026 CDS guidance expressly recognizes automation bias: the tendency to over-rely on automated suggestions, creating both commission and omission errors, particularly when time pressure reduces the practical opportunity for independent review. [2]

Peer-reviewed healthcare research has documented the same phenomenon. A systematic review found that decision support can improve overall performance while creating new error pathways when users over-rely on automation. Electronic-prescribing experiments have demonstrated automation-bias errors, and later work found that cognitive load and task complexity increase vulnerability to those errors. [17][18][19]

The opposite problem also exists. Too many alerts can produce alert fatigue and high override rates, causing clinicians to ignore software warnings. Patient safety therefore cannot be reduced to the statement: "There was a human in the loop." [20]

8. An available override is not necessarily a meaningful override

A software system can technically contain an override button while still discouraging independent judgment. If overriding software requires repeated warnings, supervisor approval, burdensome documentation, punitive review, or other institutional barriers, the professional's practical ability to exercise independent judgment may be reduced. [2][17]

FDA should distinguish between an override that merely exists and one that is usable. Meaningful oversight requires sufficient information, time, authority, and access to patient-specific context so the responsible professional can understand the output, disagree when appropriate, and prevent the computerized recommendation from automatically becoming the final action. [2]

9. Population-level risk is not an individualized treatment command

Healthcare algorithms are frequently developed using data from populations. Those models can be useful, but performance in a population does not mean that every output is correct for every individual. Predictive models can differ in calibration, generalizability, subgroup performance, threshold behavior, and clinical usefulness when applied outside the population or environment in which they were developed.[21]

Independent evaluation of a widely implemented proprietary sepsis model in 2021 found substantially weaker performance in a new health system. A newer locally trained model evaluated in 2026 performed better but still showed false positives, imperfect calibration, and meaningful differences depending on sepsis definition and clinical setting.[22][23]

A population-derived probability should remain evidence about risk rather than automatically becoming a treatment command for an individual patient. The transition from "this patient resembles patterns associated with increased risk" to "therefore this patient's treatment must be restricted" is not merely a mathematical step; it is a clinical decision.

10. Patients need a meaningful way to correct factual errors and provide context

Patients should not be trapped inside an algorithmic conclusion that depends on inaccurate information they cannot correct. HIPAA already provides an important federal policy precedent: under 45 C.F.R. § 164.526, patients may request amendment of protected health information in a designated record set, subject to defined limits.[24][25]

If a requested amendment is denied, the patient may submit a statement of disagreement. P.A.R.T. USA recommends that any patient correction, contextual explanation, or statement of disagreement remain clearly associated with the relevant record even when a healthcare professional, institution, payer, or other entity disagrees with the patient's position.[24][25]

The purpose is not to allow patients to rewrite professional medical opinions. The purpose is to prevent disputed information from circulating through healthcare systems as though no disagreement exists.

11. Patients should be alerted when material health information changes

Patients are the people whose bodies, diagnoses, medications, treatment histories, and healthcare outcomes are represented in these systems. They should not be passive outsiders to their own electronic medical information.

For consequential healthcare information, P.A.R.T. USA recommends that patients receive reasonable notice when material information in their health record is added, changed, corrected, or updated when that information may affect computerized clinical decision support, risk scoring, treatment, medication access, or other significant care.

A patient cannot help correct an error they never know exists. Because healthcare software is only as reliable as the information supplied to it, greater patient visibility into material record changes is itself a patient-safety safeguard.

12. Software-generated information should carry provenance

For consequential patient-specific outputs, healthcare professionals should be able to distinguish among patient-entered information, clinician-entered EHR information, pharmacy dispensing information, claims information, externally imported data, inferred variables, calculated values, and model-generated conclusions.[3]

A diagnosis recorded by a treating specialist is not the same thing as an inference generated from claims data. A patient-reported fact is not the same thing as a model-predicted characteristic. A medication dispensing record is not necessarily the same thing as a documented treatment indication.

13. The greater the consequence, the stronger the safeguard should be

Not all healthcare software presents the same level of risk. A scheduling program and a program whose output materially influences whether a patient receives medication should not be treated as equivalent simply because both are software. [1]

As the potential consequence of a computerized output increases, the safeguards surrounding that output should increase as well. Software capable of materially influencing diagnosis, medication access, dispensing, tapering, treatment denial, coverage, risk classification, referral, or other significant patient care should require stronger transparency, input verification, meaningful independent review, correction mechanisms, patient notification, and genuine override authority. [1][2][3]

14. False positives and false negatives are patient-safety outcomes

False positives and false negatives are not merely model statistics. In a treatment-risk system, a false positive can contribute to unnecessary treatment restriction, stigma, delayed care, undertreated pain, refusal to dispense, unnecessary tapering, or loss of beneficial therapy. A false negative can create false reassurance, missed monitoring, and missed opportunities for intervention. [2][13][23]

FDA should encourage evaluation of downstream patient consequences of false classifications, not merely overall model performance.

15. Repeated disagreement must trigger review - and systems should be reviewed periodically

Repeated disagreement between qualified healthcare professionals and a software system should itself be treated as a potential patient-safety signal. When clinicians repeatedly override the same recommendation, flag, score, or alert, the system should not merely record those overrides and continue unchanged. Those patterns should trigger formal review. [3]

P.A.R.T. USA recommends monitoring override frequency, reasons for overrides, false-positive and false-negative patterns, subgroup differences, patient complaints, treatment refusals or delays following software outputs, and recurring situations where clinicians conclude that software failed to capture relevant circumstances. [3]

In addition to continuous monitoring, consequential healthcare decision-support systems should undergo a comprehensive periodic review at least every two years, and sooner when significant safety signals, model changes, new evidence, unexplained disagreement patterns, or material changes in clinical practice occur. This two-year interval is P.A.R.T. USA's proposed best-practice standard. [3][4]

16. Software is ultimately a human creation

Even excellent software remains dependent upon its data, assumptions, validation, architecture, design, programming, updates, and the way human beings implement and interact with it. [4]

A computer software program is only as good as the human decisions behind its creation, programming, testing, implementation, and continuing oversight. Artificial intelligence does not eliminate this reality; it is developed, trained, configured, constrained, monitored, and deployed through human choices. [4]

P.A.R.T. USA Recommendations to FDA

1. **No dispositive computerized rule for consequential individual care.** A numerical threshold, score, flag, or recommendation should not automatically require medication denial, tapering, non-dispensing, diagnosis, referral refusal, treatment denial, or another consequential action without meaningful professional review.
2. **Independent human clinical judgment must remain meaningful.** Human presence alone is insufficient. The responsible professional must have information, time, and authority sufficient to evaluate the recommendation independently.
3. **Consequential software should disclose its material inputs and basis.** Users should be able to identify the patient information and assumptions materially influencing the output.
4. **Missing, corrupted, stale, or questionable data should be visible.** Software should not silently treat missing context as if it were a confirmed negative or adverse fact.
5. **Validation must match the intended use.** Validation for association, prediction, or screening should not automatically be treated as validation for denial, tapering, dispensing refusal, or another intervention.
6. **Professionals must have meaningful override authority.** An override should be practically usable, not merely technically available.
7. **Patients need a correction and context pathway.** Materially inaccurate factual inputs should be correctable, and relevant disputed or missing context should be capable of accompanying consequential outputs.
8. **Patient corrections and disagreements should remain visible.** A correction, contextual statement, or disagreement should remain associated with the relevant record even when a healthcare professional or institution disagrees.
9. **Patients should receive notice of material record changes.** Reasonable alerts should notify patients when material information is added, changed, corrected, or updated in ways that may influence consequential software-assisted care.
10. **Population-level risk estimates must not automatically become individualized treatment commands.**
11. **MME must remain a tool, never a rigid administrative threshold.** CDC itself warns against rigid dosage thresholds and inflexible implementation.
12. **Scores must be treated as general information requiring investigation.** They should trigger questions, verification, and individualized assessment - not be relied upon as absolutes.
13. **Automation bias and alert fatigue should be part of real-world safety evaluation.**
14. **Repeated clinician-software disagreement should trigger formal review.**
15. **Consequential systems should undergo comprehensive review at least every two years.** Earlier review should occur when safety signals, model changes, new evidence, or material changes in clinical practice arise.
16. **Safeguards should increase as potential harm increases.**
17. **Post-deployment monitoring should evaluate actual patient consequences.** This should include treatment denial, medication access, dispensing, tapering, delays, stigma, false classifications, patient complaints, and override behavior.
18. **Software should assist clinical questioning rather than terminate it.**

Conclusion: The Human Condition Must Remain the Final Authority

P.A.R.T. USA recognizes the enormous potential of healthcare technology. Computers can calculate, organize, compare, identify patterns, warn, recommend, and help clinicians see information that might otherwise be missed.

But medical care is delivered to individual human beings, not population averages. A computer does not automatically know why a particular treatment was chosen. It does not know that a diagnosis was entered

incorrectly unless someone identifies the error. It does not necessarily know that a patient could not afford a preferred medication. It does not automatically understand that several treatments failed before the current therapy succeeded. It does not know what information was never entered. And it cannot independently ask the next question unless it has been designed and supplied with the information to do so.

Even excellent software remains dependent upon its data, assumptions, validation, design, human programming, implementation, and continuing oversight.

> **The computer may calculate.**
> **The computer may organize.**
> **The computer may identify risk.**
> **The computer may recommend.**
> **But the computer must never become the unquestionable final authority over an individual patient's care. The human condition - the patient's actual circumstances, the clinician's judgment, the ability to ask another question, recognize an exception, challenge an assumption, and correct an error - must remain the final authority in individualized medical care.**
>
> **Technology should assist the human decision-maker. It should never replace the human condition as the final authority in individualized medical care.**

Respectfully submitted,

Ingrid "Kitty" Trimm
Founder / Executive Director
P.A.R.T. USA
Pain Awareness - Right to Treatment
Email: info@partusa.org
Phone: 833-401-7278
Website: https://partusa.org

Sources and References

[1] 21 U.S.C. § 360j(o) - Regulation of medical and certain decision support software https://www.law.cornell.edu/uscode/text/21/360j

[2] U.S. Food and Drug Administration. Clinical Decision Support Software: Guidance for Industry and Food and Drug Administration Staff. January 29, 2026. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-decision-support-software

[3] Office of the National Coordinator for Health Information Technology. Decision Support Interventions - §170.315(b)(11). https://healthit.gov/test-method/decision-support-interventions/

[4] National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0). https://airc.nist.gov/airmf-resources/airmf/3-sec-characteristics/

[5] U.S. Food and Drug Administration. Reports on Non-Device Software Functions. https://www.fda.gov/about-fda/cdrh-reports/reports-non-device-software-functions

[6] Centers for Disease Control and Prevention. CDC Clinical Practice Guideline for Prescribing Opioids for Pain - United States, 2022. https://www.cdc.gov/mmwr/volumes/71/rr/rr7103a1.htm

[7] Centers for Disease Control and Prevention. Healthcare Administrators: Applying the Guideline. https://www.cdc.gov/overdose-prevention/hcp/clinical-guidance/healthcare-admin-applying-guidelines.html

[8] Centers for Disease Control and Prevention. Prescription Drug Monitoring Programs (PDMPs). https://www.cdc.gov/overdose-prevention/hcp/clinical-guidance/prescription-drug-monitoring-programs.html

[9] U.S. Food and Drug Administration. Digital Health Policy Navigator - Step 6: Clinical Decision Support. https://www.fda.gov/medical-devices/digital-health-center-excellence/step-6-software-function-intended-provide-clinical-decision-support

[10] U.S. Department of Veterans Affairs. Buprenorphine for Chronic Pain. Veterans Health Library. https://www.veteranshealthlibrary.va.gov/DiseasesConditions/ChronicPain/142%2C41729_VA

[11] Picco L, et al. How prescription drug monitoring programs influence clinical decision-making: A mixed methods systematic review and meta-analysis. 2021. https://pubmed.ncbi.nlm.nih.gov/34600255/

[12] Haines S, Lam A, Savic M, Carter A. Patient experiences of prescription drug monitoring programs: a qualitative analysis from an Australian pharmaceutical helpline. 2022. https://pubmed.ncbi.nlm.nih.gov/36067724/

[13] McElfresh DC, et al. A call for better validation of opioid overdose risk algorithms. Journal of the American Medical Informatics Association. 2023;30(10):1741-1746. https://academic.oup.com/jamia/article/30/10/1741/7222347

[14] Cochran G, et al. Validation and threshold identification of a prescription drug monitoring program clinical opioid risk metric with the WHO ASSIST. Drug and Alcohol Dependence. 2021. https://pubmed.ncbi.nlm.nih.gov/34610516/

[15] Bamboo Health. NarxCare and Patients. https://bamboohealth.com/narxcare-and-patients/

[16] Bamboo Health. ORS Overview: Information for Healthcare Professionals. Updated March 20, 2026. https://narxcare.zendesk.com/hc/en-us/articles/18546740066835-ORS-Overview-Information-for-Healthcare-Professionals-HCPs

[17] Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. JAMIA. 2012. https://pmc.ncbi.nlm.nih.gov/articles/PMC3240751/

[18] Lyell D, et al. Automation bias in electronic prescribing. BMC Medical Informatics and Decision Making. 2017. https://pubmed.ncbi.nlm.nih.gov/28302112/

[19] Lyell D, Magrabi F, Coiera E. The Effect of Cognitive Load and Task Complexity on Automation Bias in Electronic Prescribing. Human Factors. 2018. https://pubmed.ncbi.nlm.nih.gov/29939764/

[20] Nanji KC, et al. Medication-related clinical decision-support alert override study. 2018. https://pubmed.ncbi.nlm.nih.gov/29092059/

[21] Van Calster B, et al. Calibration of risk prediction models: impact on decision-analytic performance. 2015. https://pubmed.ncbi.nlm.nih.gov/25155798/

[22] Wong A, et al. External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients. JAMA Internal Medicine. 2021. https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/2781307

[23] Dutta S, et al. Performance of a Sepsis Prediction Model Across Different Sepsis Definitions. JAMA Network Open. 2026. https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2847396

[24] U.S. Department of Health and Human Services. Your Medical Records. https://www.hhs.gov/hipaa/for-individuals/medical-records/index.html

[25] 45 C.F.R. § 164.526 - Amendment of protected health information. https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-C/part-164/subpart-E/section-164.526

Source note: This document distinguishes between existing legal/regulatory requirements and P.A.R.T. USA policy recommendations. The proposed two-year comprehensive review interval and patient alert mechanism are recommendations advanced by P.A.R.T. USA; they are not presented as current universal FDA mandates.