To: U.S. Department of Health and Human Services
Re: Specific Recommendations Concerning Healthcare and AI
From: Public Citizen

The Biden Administration issued a comprehensive Executive Action on Artificial Intelligence (AI) which included a number of directives for agencies dealing with healthcare, public health and human services. Below we share our recommendations.

EO Directive: Establish an HHS AI Task Force – within 90 days of the AI EO – that is to develop a strategic plan on responsible deployment and use of AI and AI-enabled technologies in the health and human services sector – within 1 year of the establishment of the HHS AI Task Force – on the following areas:

Recommendations**:** In addition to subject matter and technical experts, we would recommend regular and consistent engagement with civil society groups, patient advocates, and representatives of marginalized populations. This should include regular updates from the taskforce to stakeholders, opportunities for stakeholders to provide comments (e.g., via requests for information (RFIs) and through rulemaking) in writing, and recurring gatherings of stakeholders to provide updates and elicit feedback.

EO Taskforce Area Sub-Directive: development, maintenance, and use of predictive and generative AI-enabled technologies in healthcare delivery and financing — including quality measurement, performance improvement, program integrity, benefits administration, and patient experience — taking into account considerations such as appropriate human oversight of the application of AI-generated output;

Recommendations: Requirements around human oversight of AI are particularly important and should represent a central protection in all aspects of AI development and use in health care. Testing should take place in a way that limits potential harms by starting with simulations or closely overseen demonstrations. The taskforce should identify a baseline of relevant metrics and appropriate methodologies to track and also test the effects of using AI under controlled and constrained circumstances. This will be crucial to limit harms from too rapid a rollout.

Whenever an AI system is used to make health decisions that may have an impact on a patient, the patient and their physician should have the right to an understandable explanation of the decision, the right to request human review, and the right to have the decision appealed to a human.

EO Taskforce Area Sub-Directive: long-term safety and real-world performance monitoring of AI-enabled technologies in the health and human services sector, including clinically relevant or significant modifications and performance across population groups, with a means to communicate product updates to regulators, developers, and users.

Recommendations: It will be important for this information to be available to the public, particularly researchers and civil society groups, in a searchable, sortable, and downloadable

format. The bodies charged with this monitoring will need to be accountable to the public and should provide as much transparency around the data as possible. These bodies will also likely need statutory access to provider and institution reported data in a timely manner. This will likely mean the need for enforcement authority, potentially including escalating enforcement actions, such as fines, Medicare payment freezes, or even termination of a provider agreement in the event providers and institutions are not providing required data in a timely manner.

On the provider side, the taskforce should pay special attention to consumer healthcare AI technologies, including unregistered tools and apps – such as chatbots with explicit and implicit therapeutic claims – and generalized tools that consumers may use for therapeutic purposes or medical advice. The taskforce should evaluate reliance on these tools, how they connect to the formal health care system, and evidence of their efficacy, as well as how the proprietors of these tools can and should measure impact.

EO Taskforce Area Sub-Directive: incorporation of equity principles in AI-enabled technologies used in the health and human services sector, using disaggregated data on affected populations and representative population data sets when developing new models, monitoring algorithmic performance against discrimination and bias in existing models, and helping to identify and mitigate discrimination and bias in current systems.

Recommendations: Strong equity principles must underlie this effort and must seek to improve equity across a number of factors, including (but not limited to) race, ethnicity, sex, language, sexual orientation, disability status, gender identity, age, and socioeconomic status. Companies and providers must be given clear guidance on how equity principles are to be incorporated into their AI-enabled technologies and what data are expected to be collected and reported in a timely manner (as well as the penalties for failure to report). Monitoring will be important, but options for enhanced scrutiny, including targeted audits and accountability and enforcement mechanisms, will also be important to ensure proper protections for patients. Baseline data and data over time will need to be made available to regulators, researchers, stakeholders, and the public to allow a better understanding of what role AI-enabled technologies are playing in health equity.

EO Taskforce Area Sub-Directive: incorporation of safety, privacy, and security standards into the software-development lifecycle for protection of personally identifiable information, including measures to address AI-enhanced cybersecurity threats in the health and human services sector.

Recommendations: Clear HHS guidance and, likely, rulemaking will be important in this area, both for creating standards to limit potential harms and to create accountability when actors are not taking the necessary steps to protect such vitally important patient data. Rulemaking should include what protections are required for sharing, storing, and anonymizing data for AI-enabled devices above and beyond current data requirements; enforcement efforts and penalties for failure to meet these standards; and mechanisms through which patients can choose to have their data not be included in an AI database and for their data to be deleted from such a

database. HHS rules and guidance should establish their applicability to general AI tools or unregistered tools that make health-related claims.

EO Taskforce Area Sub-Directive: development, maintenance, and availability of documentation to help users determine appropriate and safe uses of AI in local settings in the health and human services sector;

Recommendations: Engaging providers and improving their understanding of both the potential benefits and risks of AI will be necessary. There will likely be a need for multiple forms of documentation, both broad and specific, given the broad range of potential uses of AI within the medical field. Specificity should include the risks both to patients and to providers in the event that an AI enabled system provides incorrect information which leads to an adverse event. Regional HHS offices should be marshalled to help specify this documentation to meet regional needs and to work with providers and institutions to understand additional needs for documentation or guidance.

EO Taskforce Area Sub-Directive: work to be done with State, local, Tribal, and territorial (SLTT) health and human services agencies to advance positive use cases and best practices for use of AI in local settings; and

Recommendations: HHS should play a leading role in pulling together relevant stakeholders given how fragmented our health care system is, but how central HHS could be on this. Given the potential risks of the use of the use of AI in health care, in addition to positive use cases and best practices, HHS should also provide clear-eyed assessments of the risks of the use of AI, the need for strong data protections, and examples of adverse events of AI and steps to take to mitigate the risks of such events. A clear and consistent schedule of convenings and publications would help SLTT stakeholders remain abreast of developments and have multiple opportunities to engage.

EO Taskforce Area Sub-Directive: identification of uses of AI to promote workplace efficiency and satisfaction in the health and human services sector, including reducing administrative burdens.

Recommendations: While there are promising reports of the use of AI for reducing administrative burden, including developing fair and more consistent schedules for workers, booking of operating room suites, and filling gaps in providers' schedules, it will be important for HHS to help identify best practices and recommendations for implementation. However, such use of AI is not without risks and mitigation efforts and missteps should also be a part of any documentation. Further, such AI uses should continue to have a human in the loop to help ensure that decisions are being made in an expected and appropriate manner. The charge of efficiency should not come at the expense of patients or providers, as we have seen with the use of AI decision-making systems by Medicare Advantage plans.

EO Directive: HHS shall develop a strategy on maintaining appropriate levels of quality in AI-enabled technologies the in health and human services sector – within 180 days of the AI EO

Recommendations: Measuring appropriate quality across the many possible use of AI in health care will be challenging and require significant resources and broad authority. Identifying any additional authority the agency needs on this will be important, given the broad remit of this directive. Appropriate quality measures should include things that are relevant to patients, providers, and institutions. Patient quality measures should include both objective measures related to improvements in measurable outcomes as well as subjective standards, such as having patients provide a grade for the services they received and assessments of changes in their wellbeing (quasi-experimental studies comparing the experience of patients who interacted with AI-enabled technology with those who did not within a similar setting seem like an important source of potential data). Similarly, for providers and institutions, objective measures that indicate the extent to which AI-enabled technologies facilitated affected provider experience, efficiency of clinical practices, or stabilized provider hours would be useful. Subjective measures, where providers assess the extent to which specific AI-enabled technologies improved their ability to practice medicine or reduced feelings of burnout could help inform necessary changes and further improvements.

EO Directive: HHS shall consider appropriate actions to advance understanding of and compliance with Federal nondiscrimination laws by providers that receive federal financial assistance, as well as how those laws relate to AI – within 180 days of the AI EO

Recommendations: This is a particular area of concern given the denials and delays experienced by Medicare Advantage enrollees, so we appreciate the steps that the Biden Administration has already taken on curtailing the use of AI-enabled tools for denials of care by Medicare Advantage plans. However, given the potential for widespread use of AI across health care, it will be crucial to anticipate potential areas of concern and respond quickly to examples of AI-enabled tools exacerbating discrimination in health care. It will be important to collect relevant data and have sufficient resources for the analysis and the reporting of that data by relevant demographic and other factors where discrimination may be evident. HHS should also take action to help address concerns about the potential for AI to be trained on data sources that contain biased samples or inherent limitations that may exacerbate discrimination in health care by setting standards for the use of databases for training AI and for creating requirements around transparency of the databases on which AI enabled technologies are being trained. Much of the work of the Biden Administration on advancing racial justice and gender equality can serve as guidance for ensuring that AI-enabled technology in the health care space is being used appropriately and assessed properly.

EO Directive: HHS shall establish an AI safety program – within 365 days of the AI EO – that establishes a common framework for approaches to identifying and capturing clinical errors resulting from AI deployed in healthcare settings as well as specifications for a central tracking repository for associated incidents that cause harm, including through bias or discrimination, to patients, caregivers, or other parties.

Recommendations: The gathering and dissemination of this information will be particularly important for researchers, patient advocates, and policymakers. A consistent schedule of dissemination of such data will allow researchers to engage with it in an iterative way, improving both our understanding of what the data are telling us and likely leading to better use of AI-enabled technology. The granularity of data collected versus reported will also be important. HHS will need to be collecting granular data, even if the data being provided or reported on is less granular (to potentially protect patient populations). It seems that the missing piece may be some tools for enforcement in this area, based on the data being received, given the consistent data failures of companies charged with reporting data to the federal government, including Medicare Advantage insurers. Without sufficient penalties to ensure compliance with data reporting requirements, companies may prefer to only report limited data (which may be biased) or not report at all.

EO Safety Program Sub-Directive: analyzes captured data and generated evidence to develop, wherever appropriate, recommendations, best practices, or other informal guidelines aimed at avoiding these harms; and disseminates those recommendations, best practices, or other informal guidance to appropriate stakeholders, including healthcare providers.

Recommendations: Having HHS also undertake significant analysis, both using researchers and potentially AI (e.g., flagging potentially concerning patterns in the data) will be necessary. While the EO recommends informal guidelines, which will play a role, this is likely an area where formal guidance, rulemaking, and even statutory changes will be necessary to protect patients, providers, and the general public from worst practices of AI-enabled technologies, including potential fraud, waste, and abuse. Engaging with stakeholders through regular processes would help contextualize the results being identified in the data and what steps have been found to be effective in mitigating risks and harms.

EO Directive: HHS shall develop a strategy for regulating the use of AI or AI-enabled tools in the drug-development process – within 365 days of the AI EO – to include: define the objectives, goals, and high-level principles required for appropriate regulation throughout each phase of drug development; identify areas where future rulemaking, guidance, or additional statutory authority may be necessary to implement such a regulatory system; identify the existing budget, resources, personnel, and potential for new public/private partnerships necessary for such a regulatory system; and consider risks identified by the actions undertaken to implement section 4 of this order.

Recommendations: World Health Organization experts have identified key benefits and risks of using AI for pharmaceutical R&D and delivery that should be reflected in the HHS strategy, including:

- AI and AI-enabled tools used in drug development must help meet public health needs, including in rare and traditionally neglected disease areas, rather than only benefit private interests and pharmaceutical company profits and patients with the greatest financial means and disease areas with the most available data.

- Exclusive control of AI algorithms should be avoided to maximize its utility in promoting research, especially when developed with public support. Patent and trade secret barriers will likely need to be overcome to accomplish this, through existing and additional statutory authority.
- Use of AI by drug manufacturers in pricing and marketing strategies must be fully transparent to the public and regulators, especially when intended to influence prescribing behavior, whether they target patients or physicians.
- Privacy of patient data used in AI-based drug development must be preserved. Legally collected patient data used in AI-based drug development should be made openly available to promote research.

Additionally, medicines developed with AI and AI-enabled tools in government laboratories or in whole or part with federal funding should be made available to the public on reasonable terms, including fair prices that reflect the public's investment.