TECHNOLOGY ASSESSMENT

ARTIFICIAL INTELLIGENCE IN HEALTH CARE

Benefits and Challenges of Machine Learning in Drug Development

What GAO Found

Machine learning—a field of artificial intelligence (AI) in which software learns from data to perform a task—is already used in drug development and holds the potential to transform the field, according to stakeholders such as agency officials, industry representatives, and academic researchers. Machine learning is used throughout the drug development process and could increase its efficiency and effectiveness, decreasing the time and cost required to bring new drugs to market. These improvements could save lives and reduce suffering by getting drugs to patients in need more quickly, and could allow researchers to invest more resources in areas such as rare or orphan diseases.

Machine learning could accelerate drug development

This set of technologies could screen more chemical compounds and zero in on promising drug candidates in less time than the current process.

Examples of machine learning in the early steps of drug development include:
- **Drug Discovery**: Researchers are identifying new drug targets, screening known compounds for new therapeutic applications, and designing new drug candidates, among other applications.
- **Preclinical Research**: Researchers are augmenting preclinical testing and predicting toxicity before testing potential drugs in humans.
- **Clinical Trials**: Researchers are beginning to improve clinical trial design, a point where many drug candidates fail. Their efforts include applying machine learning to patient selection, recruitment, and stratification.

GAO identified several challenges that hinder the adoption and impact of machine learning in drug development. Gaps in research in biology, chemistry, and machine learning limit the understanding of and impact in this area. A shortage of high-quality data, which are required for machine learning to be effective, is another challenge. Accessing and sharing these data is also difficult, due to costs, legal issues, and a lack of incentives for sharing. Furthermore, a low supply of skilled and interdisciplinary workers creates hiring and retention challenges for drug companies. Lastly, uncertainty about potential regulation of machine learning used in drug development may limit investment in this field.
GAO developed six policy options in response to these challenges. Five policy options are centered around research, data access, standardization, human capital, and regulatory certainty. The last is the status quo, whereby policymakers—federal agencies, state and local governments, academic and research institutions, and industry, among others—would not intervene with current efforts. See below for details of the policy options and relevant opportunities and considerations.

Policy Options to Address Challenges to the Use of Machine Learning in Drug Development

<table>
<thead>
<tr>
<th>Policy Option</th>
<th>Opportunities</th>
<th>Considerations</th>
</tr>
</thead>
</table>
| **Research** (report page 27) | Policymakers could promote basic research to generate more and better data and improve understanding of machine learning in drug development. | • Could result in increased scientific and technological output by solving previously challenging problems.  
• Could result in the generation of additional high-quality, machine readable data. |
| **Data Access** (report page 28) | Policymakers could create mechanisms or incentives for increased sharing of high-quality data held by public or private actors, while also ensuring protection of patient data. | • Could shorten the length of the drug development process and reduce costs.  
• Could help companies identify unsuccessful drug candidates sooner, conserving resources. |
| **Standardization** (report page 29) | Policymakers could collaborate with relevant stakeholders to establish uniform standards for data and algorithms. | • Could improve interoperability by more easily allowing researchers to combine different data sets.  
• Could help efforts to ensure algorithms remain explainable and transparent, as well as aid data scientists with benchmarking. |
| **Human Capital** (report page 30) | Policymakers could create opportunities for more public and private sector workers to develop appropriate skills. | • Could provide a larger pool of skilled workers for agencies, companies, and other research organizations, allowing them to better leverage advances in the use of machine learning in drug development.  
• Interdisciplinary teamwork could improve as workers with different backgrounds learn to better communicate with one another. |
| **Regulatory Certainty** (report page 31) | Policymakers could collaborate with relevant stakeholders to develop a clear and consistent message regarding regulation of machine learning in drug development. | • Could help increase the level of public discourse surrounding the technology and allow regulators and the public to better understand its use.  
• Drug companies could better leverage the technology if they have increased certainty surrounding how, if at all, regulators will review or approve the machine learning algorithms used in drug development. |
| **Status Quo** (report page 32) | Policymakers could maintain the status quo (i.e., allow current efforts to proceed without intervention). | • Challenges may be resolved through current efforts.  
• Companies are already using machine learning and may not need action from policymakers to continue expanding its use. |

Source: GAO.