Technological Forecasting & Social Change 198 (2024) 122940

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

Did the COVID-19 pandemic propel usage of AI in pharmaceutical innovation? New evidence from patenting data☆

Sawan Rathi [a], Adrija Majumdar [b], Chirantan Chatterjee [c,*]

[a] Science Policy Research Unit, Business School, University of Sussex, UK & IIM Ahmedabad, India
[b] Information Systems Area, IIM Ahmedabad, India
[c] Science Policy Research Unit, Business School, University of Sussex, UK & Hoover Institution, Stanford University, USA

ARTICLE INFO

JEL codes:
O31
O32
O33
I10

Keywords:
Innovation management
AI
Pharmaceutical industry
Patents
Pandemic

ABSTRACT

It is now much discussed that Artificial Intelligence (AI) as a General-Purpose Technology (GPT) can resolve the efficiency problems of industries, including in pharmaceutical markets where productivity challenges continue in costs and time for new drug discovery. But did the COVID-19 pandemic inadvertently accelerate the pace of AI adoption in pharmaceutical innovation? We answer this question using novel data on pharmaceutical patents. We use two different databases to analyze abstracts of pharmaceutical patents applied in the USA. Topic modeling was used to identify patents with technical artifacts and classify them as treated group AI-adopting patents. An AI dictionary is used to match AI-related keywords in the patent abstracts. Subsequently, using a difference-in-differences research design we observe that both presence and count of AI keywords in pharmaceutical patents have increased with pandemic. An increase in AI is also related to reduced time taken from application to publication of a patent suggesting innovation efficiencies in the industry. Finally, we find that results are driven by firms that have already built AI capability in the past. Our results remain consistent with various robustness checks, and we conclude by discussing managerial and policy implications of our findings.

> "COVID-19 has fast-tracked the usage of AI in disease identification, drug discovery, clinical trials, and predictive forecasting" – GlaxoSmithKline[1]

1. Introduction

There has been a rise in the investigation and use of AI and other data analytic tools in multiple areas since the outbreak of the pandemic (Sipior, 2020).[2] The pandemic has affected global health, individuals, and firms differently. From an individual's perspective, it is important to get timely treatment and effective vaccines in times of high uncertainty and evolving virus variants (Muguerza, 2020).[3] From a firm's perspective, sustainable innovation has become an important aspect for survival with rising global costs post-pandemic (Kramme, 2021; Lee and Trimi, 2021). Amidst all this, COVID-19 has put a spotlight on AI as a GPT, which is now seen as to be playing an important role in providing quick, sustainable and affordable solutions in the biotech field (Block, 2020).[4]

AI as a GPT has received increasing interest in recent work in the economics of technological change (Agrawal et al., 2022; Bianchini

☆ We thank Ritu Agrawal, Pankaj Setia, and Manisha Rathi for their helpful comments and suggestions. We also thank participants of conference at EU-SPRI Annual Conference at University of Sussex, Summer School on Data & Algorithms for STI studies at KU Leuven, Belgium and International Conference of Digital Organizations (ICODO'23) at IIM Ahmedabad for useful discussions. Chatterjee acknowledges support from the Hoover Institution Visiting Fellowship at Stanford University and University of Sussex Business School research grants. He also acknowledges insights gained from his consulting work on AI and Lifesciences with Marlabs Inc. and Compass Lexecon.

* Corresponding author.
E-mail addresses: s.rathi@sussex.ac.uk (S. Rathi), adrija@iima.ac.in (A. Majumdar), c.chatterjee@sussex.ac.uk (C. Chatterjee).

[1] https://www.forbes.com/sites/cognitiveworld/2020/12/26/the-increasing-use-of-ai-in-the-pharmaceuticalindustry/?sh=3214767a4c01.

[2] A recent study by Kearney revealed that 68 % of industry leaders in the healthcare sector see AI and advanced analytics as major value drivers. https://blog.marketresearch.com/covid-19-accelerates-use-of-ai-in-pharma.

[3] https://pharmaphorum.com/views-analysis-digital/ai-infiltration-of-pharma-how-covid-19-acceleratedchange/.

[4] https://www.genengnews.com/gen-edge/covid-19-puts-spotlight-on-artificial-intelligence.

https://doi.org/10.1016/j.techfore.2023.122940
Received 13 January 2023; Received in revised form 11 September 2023; Accepted 19 October 2023
Available online 7 November 2023

et al., 2022; Cockburn et al., 2018; Trajtenberg, 2018). Scholars have studied the antecedents to how the adoption of AI by firms and industries resulted in product innovation and changed employee profiles (Haefner et al., 2021; Mikalef and Gupta, 2021; Rammer et al., 2022; Yang, 2022). Evidence also suggests that size of firm matters herein (Davenport and Ronanki, 2018; Grimaldi et al., 2013), as also human absorptive capacity (Agrawal et al., 2019; Cao et al., 2021; Mikalef and Gupta, 2021) along with regulatory environment (Hammadou et al., 2014; Porter and Stern, 2001). However, there are limited studies on how necessity of innovation management can propel the usage of AI.

In this study, we investigate how AI adoption increased in pharmaceutical innovation during the COVID-19 pandemic. Our context is important from a societal welfare perspective because past work has indicated that the pharmaceutical industry is going through a productivity crisis (Pammolli et al., 2011). It takes 12 years and an average of almost $2 billion to bring a drug to market (Abramo et al., 2011; Burki, 2019; DiMasi et al., 2016). This translates into high prices of drugs like the recently approved most expensive medicine Hemgenix, which is priced at $3.5 million a dose to treat Haemophilia B.[5] It would thus be important to know if COVID-19 unintendedly accelerated the adoption of AI in the pharmaceutical innovation process, potentially rendering better productivity that can ultimately have positive societal welfare consequences with reduced market power exercised by firms (Bloch and Metcalfe, 2018; Schumpeter, 1942).

A solution to the pharmaceutical productivity crises has been to compress drug development timelines (Kummar et al., 2007), and AI has been argued to be useful in achieving that (Shanbhogue et al., 2021). AI can help the pharmaceutical sector in different dimensions, such as developing new drugs, tackling diseases previously thought to be difficult to manage, interpreting clinical data, and finding appropriate patients for clinical trials (Zielinski, 2021).[6] As per Babu et al. (2021), AI is supposed to revolutionize how drugs are discovered, along with amplifying several tasks in pharmacy and overall medical care.

Drug discovery is a complex process that can be divided into four stages: (i) target selection and validation; (ii) compound screening and lead optimization; (iii) preclinical studies; and (iv) clinical trials. Major contribution of AI, as of now, it seems, has been in stage 2 as it reduces cost and time when compound screening and lead optimization are done (Smalley, 2017). A traditional screening library contains around one million compounds, and each compound typically costs $50 –$100 (Green, 2017). Thus, an initial screening process can cost millions plus several months of work. Subsequent lead optimization can take several years to pinpoint preclinical drug candidates. With AI, a virtual compound library of billion molecules can be screened in a few days (Chan et al., 2019). Additionally, an AI-based computational pipeline might only take a few months to identify preclinical candidates.[7]

AI growth is still in the early phase, and data post COVID-19 is too small to see the effects at all stages of drug discovery. Much work here might also be tacit knowledge and not visible through patenting data because firms might be holding back from filing patents in important complex steps where they use AI – purely for competitive intelligence purposes vis-a-vis market peers. That said, some evidence still exists that AI has brought significant change in drug discovery. For example, an AI company, Recursion Pharmaceuticals, in collaboration with Takeda Pharmaceutical Ltd., announced breakthrough results in identifying preclinical compounds for rare diseases recently. Interestingly, they could do it in 1.5 years while ideally the traditional preclinical drug discovery takes a decade.[8] Similarly, Merck has successfully used Deep Learning algorithms for predicting native protein folding which can now be achieved within a few days. This process used to take several years (Bada, 2019).

Though AI is considered an important tool in business innovation, the adoption rate until the recent past has not been inspiring (KPMG, 2019), especially in healthcare (Goldfarb and Teodoridis, 2022; Stempniak, 2022). Tradeoffs in adopting AI may include technological as well as wage inequality effects (Acemoglu and Restrepo, 2020). New work seems to suggest the importance of complementary investments required along with AI implementation (Brynjolfsson et al., 2021). There are also cultural issues to deal with as workforce composition changes with firms and industries adopting AI to enhance their innovation (Yang, 2022). Thus, there are broader managerial ramifications within and beyond the boundary of the traditional bio-pharmaceutical firm.

Several studies seem to indicate that healthcare at large and pharmaceuticals, in particular, seem to have been disrupted during COVID-19 days in their innovation organization (Block, 2020; Irving, 2020; Kramme, 2021; Lee and Trimi, 2021; Rathi et al., 2022; Tietze et al., 2020).[9] But a systematic investigation of whether the adoption is actually happening and does it translate into reduced timelines is not yet done especially using patenting data from the industry.

In this paper, we investigate this question using two important patent datasets. Using Derwent Innovation database[10] we analyze abstracts of all pharmaceutical patents filed in the U.S. from October 2019 to August 2020. We use Latent Dirichlet Allocation (LDA) based topic modeling as an unsupervised technique and identify hidden topics within the patent abstracts.[11] Next, we identify the topic (Topic 14)[12] that comprise keywords representing technical artifacts and use that to divide all patents into two groups.[13] Abstract of the patents that comprise Topic 14 in highest probability are considered as Treated Group Patents (or patents that likely use AI) and others are Control Group Patents (patents that don't use AI).

To identify the presence of AI in these patent abstracts, we use an AI dictionary comprising all artificial intelligence-related functional applications and the key phrases (See Table A1 for excerpts of AI Dictionary keywords). The main source of this dictionary is the Patentscope artificial intelligence index by World Intellectual Property Organization (WIPO).[14] We search for words in the AI dictionary in all patent abstracts in our sample (see Table A2 for an example).[15] We check how the presence and count of usage of these words changed in treated group patents as compared to control group patents during COVID-19. Using difference-in-differences we find that both the presence and count of AI

[5] https://www.bloomberg.com/news/articles/2022-11-23/world-s-most-expensive-drug-csl-hemgenixhemophilia-approved-by-fda?leadSource=uverify%20wall.

[6] https://www.digitalauthority.me/resources/artificial-intelligence-pharma/.

[7] https://www.prnewswire.com/in/news-releases/insilico-medicine-achieves-industry-first-nominatingpreclinical-candidate-discovered-by-ai-877449436.html.

[8] https://www.businesswire.com/news/home/20190107005349/en/Recursion-Announces-Options-Exercise-Takeda-Extension-AI-enabled.

[9] BlueDot, a Canadian artificial intelligence platform, detected a cluster of pneumonia cases in Wuhan before the world even knew about COVID-19.

[10] For more than 50 years, Derwent™ has been a leader in life sciences – connecting research applications, exclusive patent data, and expert IP services to support the commercial success of pharmaceutical and biotechnology companies around the world.

[11] Prior work has used topic modeling to identify disruptive technologies (Momeni and Rost, 2016) and the commercial viability of healthcare innovations (Erzurumlu and Pachamanova, 2020)

[12] As we can see in Table 1 Topic 14 comprises most of the keywords with technical artifacts. Detailed keyword cloud of Topic 14 can be seen in Fig. 2.

[13] Since AI is an advanced technology, its important to focus on patents that comprise technical artifacts and how presence of AI among these technical artifacts changes during COVID-19.

[14] https://www.wipo.int/techtrends/en/artificialintelligence/patentscope.html.

[15] We understand that the AI Dictionary keywords may have some overlap with software and business method patents. To the extent there may be some overlap, our results might be a lower bound.

keywords have significantly increased in treated group patents filed during the pandemic. To evaluate the impact that use of increased AI has had on patents, we calculate the time lapsed from patent application to publication. We find that presence of AI has interestingly speeded up the publication process of patents for pharmaceutical firms. A detailed pictorial representation of the flow of our research is shown in Fig. 1. We conduct multiple robustness and falsification checks to strengthen our claim and buttress our identification strategy.

We cross-check these results using Artificial Intelligence Patent Dataset (AIPD) (Giczy et al., 2022) from the United States Patent and Trademark Office (USPTO). AIPD dataset is generated using a machine learning (ML) approach that analyzes patent text and citations to identify patents with AI.[16] Along with the indicator that identifies whether AI is used, AIPD also provides information of the presence of 8 different components of AI in the patents (Details in Section 3.2). We find that likelihood of usage of AI along with most of its components has significantly increased during COVID-19 in pharmaceutical patents. This is followed by a significant reduction in time to publication that we now observe in both datasets.

In further investigation, we identify the firms in our sample which applied for patents with a high presence of AI and classify these firms as leader firms, while others who don't use AI classified as laggard firms. We analyze patents filed by the leader and laggard firms two years before our study period (2018–2019). We find a consistent trend of innovation capability building done by leader firms. Thus, the mechanism of emphasis on the usage of AI in pharmaceutical research during the pandemic can be associated with apriori built-up capabilities.

All these findings highlight the likely normative shift in the adoption of AI in pharmaceutical innovation, thereby enhancing the role of AI as a GPT. As per Bresnahan and Trajtenberg (1995), a GPT should possess three characteristics: pervasiveness, improvement and innovation spawning. Our results in many ways take ahead the debate of AI as GPT as we show how AI gets quickly adopted by pharmaceutical firms nudged by the pandemic shock. Ours is the first study to investigate the adoption of AI in innovation of industry, specifically in pharma, and hence it is important from a public policy perspective. We seem to find robust partial equilibrium evidence which is tested on two different datasets to inform that AI is increasingly used by pharmaceutical firms during COVID days leading to the quickening of the publication process.

The paper proceeds as follows. In Section 2, we discuss the related literature and present our research questions. In Section 3, we describe our data sources and methodology. This is followed by Section 4, where we report our findings. In Section 5, we conduct multiple robustness checks. Finally, in Section 6, we discuss policy implications and conclude.

2. Literature review & research questions

2.1. AI emerging as GPT

Prior work suggests that there are five key attributes that define a technology to be defined as emerging - radical novelty, fast growth, coherence, prominent impact, and ambiguity & uncertainty (Rotolo et al., 2015). AI confirms all these attributes of being an emerging technology (Aristodemou and Tietze, 2018; Bianchini et al., 2022). What is interesting to know is whether AI emerges as a GPT. Bresnahan and Trajtenberg (1995) argue that a GPT should spread across sectors (pervasiveness), should get better over time (improvement), and should make it easier to invent (innovation spawning). To prove its pervasiveness AI has already spread across sectors (Bresnahan, 2010). A few examples could be from noticing how in the machinery and equipment industry AI is used through virtual factories that simulate the production process and improve efficiency (Nolan, 2020). In pharmaceuticals, prior work has shown that, on implementation of AI, delay in drug development and failure at the marketing and clinical level can be reduced (Shanbhogue et al., 2021).

To demonstrate its improvement characteristic, AI is upgrading the process of R&D by extensive use of enhanced prediction algorithms and large datasets (Cockburn et al., 2018). AI is bringing forth new methods for research and invention. Deep learning methods and prediction technologies are influencing the knowledge production process by increasing the efficiency of searching relevant prior knowledge and by easing the discovery of new results (Agrawal et al., 2019; Bianchini et al., 2022). AI-based assistive tools are increasingly used for drug repurposing (Zhou et al., 2020) as also can be seen from works by MIT's Broad Institute.

AI is likely also innovation spawning as firms that have applied AI using different methods in different applications areas have obtained experience in using AI that provides them significantly higher innovation output (Rammer et al., 2022). Thus, AI, in many ways, fulfills the requirements of being pervasive, improving, and innovation-spawning for it to be characterized as a GPT. Our research contributes to this literature by showing how AI as a GPT is increasingly used in the pharmaceutical industry and is giving rise to improvement in its innovation timelines.

2.2. Innovation management in pharma

The pharmaceutical industry seems to be under serious productivity crises (Cockburn, 2006; Schuhmacher et al., 2021). The rising cost of newly approved drugs, late-stage abandonment of drug development projects, and proliferation of plausible targets resulting from advances in molecular biology are major woes of the pharmaceutical industry (Pammolli et al., 2011). Ideally, given the unique ability of humans to be creative, innovation, many argue, should be the domain of humans (Amabile, 2020). But in the contemporaneous world, innovation needs to be challenged by introducing AI and ML because of their cost advantages, higher quality, and greater efficiency than humans (Agrawal et al., 2019; Bughin et al., 2018). Finding ways to apply AI to firms' innovation processes should be of substantial interest to innovation managers in present times (Calvino et al., 2022; Dernis et al., 2023; Haefner et al., 2021).

Innovation, specifically in the pharmaceutical industry, is driven by the drug development process, which depends on compound discovery and searches through combinatoric space (Lou and Wu, 2021). In drug development, input, behavior, and output should be simultaneously controlled with the focus on radical and incremental innovation (Cardinal, 2001). AI and ML tools offer the promise of revolutionizing drug development and overcoming hurdles in the drug discovery pipeline (Malandraki-Miller and Riley, 2021).[17] AI can also play a part at the design level by having an empowering effect on knowledge translation at different levels (Dal Mas et al., 2020) and nurture open innovation (Secundo et al., 2020) which is the need of the pandemic-ridden world today.

From the literature above, we understand that pharmaceutical firms need streamlined innovation management, and AI as GPT could be a viable solution. We build upon this literature to construct our research questions for this study. We start with a core construction on whether the COVID-19-induced pandemic propelled AI usage in pharmaceutical innovation.

Research Question 1: *Did the likelihood of the presence and count of AI keywords in pharmaceutical patent abstracts increase during COVID-19?*

The healthcare ecosystem comprises various players that can be classified as regulators, providers, payers, suppliers, clinicians, and

[16] https://www.uspto.gov/ip-policy/economic-research/research-datasets/artificial-intelligence-patent-dataset.

[17] AI can help screen compounds 100 times faster than humans can by using conventional methods (Wu et al., 2020).

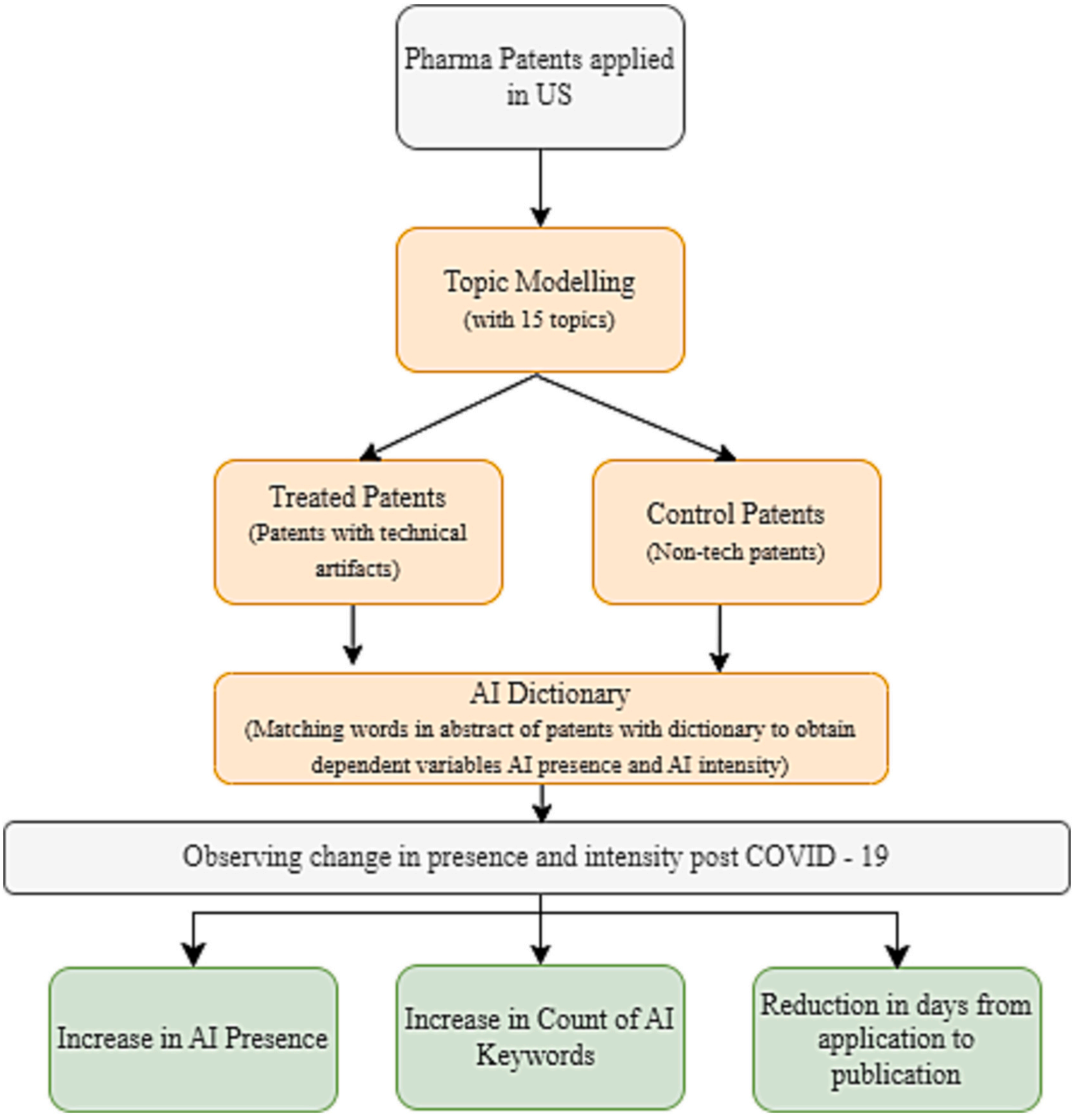

Fig. 1. Research flow diagram. In the pictorial representation, we show that we first shortlist pharmaceutical patents filed in the US from the Derwent database. We then apply Topic Modeling, which helps us bifurcate treatment and control groups of patents. Using AI dictionary we find if abstract of these patents has AI present along with its count. Eventually, we show how COVID-19 caused increase in presence and count of AI keywords on pharmaceutical patents which also led to reduction in duration of patent publication. In addition to patents from Derwent we also used AIPD dataset.

patients (Bessant et al., 2012). It is worth noting that extant literature suggests that patent examiners and the patent examination process are not homogeneous (Cockburn et al., 2002). Although the examination process of a patent office is standardized, it is imperfect in that it substantially depends on the examiners' experience, motivation, and skills (Kim and Oh, 2017). The duration between the application to publication process depends a lot on the quality of patent and US examiners tend to devote more search effort to weaker patents (Lei and Wright, 2017).

There are also instances where firms lobby for their patents to publish faster. For example, Tabakovic and Wollmann (2018) find that patent examiners grant more patents to the firms that hire them later and that much of this leniency extends to eventual employers. This strand in the literature on economics of innovation thus addresses a key concern, which is that firms actually can influence the examination process and hence time to publication.

It is also worth reflecting on the institutional structure of the patent examination process. When entering an examination, a patent application is assigned to an art unit based on the subject matter of the invention. The art unit is an administrative set of eight to fifteen patent examiners specializing in a particular technology responsible for the examination. Palangkaraya et al. (2011) find that if the patent office possesses specialization, it is less likely to misclassify an application which may result in a longer duration of examination. Building on this finding, one can hypothesize that specialized examiners may have believed that AI and related technology can lead us towards new medicines like Covid-19 vaccines more quickly, especially if tacit knowledge embedded within patents using AI is linear and simplified. This is potentially why also it's important to highlight why the AI effect in pharmaceutical patents is translating into reduced publication time. This leads to our second research question.

Research Question 2: *Was there a reduction in the duration for publishing patents comprising AI in the pharmaceutical industry during COVID-19?*

AI contributes at the resource level not just in generating and developing ideas but also helps remove barriers to innovation such as ineffective search routines and information processing constraints (Haefner et al., 2021). To realize performance gains from AI, it's important to examine how firms build on AI capability (Mikalef and Gupta, 2021). To create sustainable AI capability, organizations require a unique blend of physical, human, and organizational resources (Davenport and Ronanki, 2018), which can deliver value by differentiating it from that of competitors (Igna and Venturini, 2023; Ransbotham et al., 2018). Drawing from the resource-based theory (RBT) perspective, it's important for a firm to build up capability and resource complementarity to react in time when hit by a shock (Priem and Butler, 2001). In line with this thought, we propose that firms that had built apriori capabilities in the past were nudged in the COVID-19 era to adopt AI to change their innovation process more than those that hadn't. We summarize our proposition below formally.

Research Question 3: *Did firms that increasingly used AI during COVID-19 have built-in AI capability in past years?*

In the next section, we will empirically evaluate these research questions.

3. Data and methodology

We use the Derwent Innovation database and the AIPD dataset for our study. Owing to the time lag involved from ideation to the publication stage in a patent, we use the application date of the patent in both databases to determine whether a patent was initiated before or after the COVID-19 shock.

3.1. Derwent database

The study uses abstracts of 30,527 pharmaceutical patents filed in the US between October 2019 to August 2020 obtained from Derwent Innovation.[18][19] We do topic modeling using LDA on the abstracts of the pharmaceutical patents to render constructs and conceptual relationships from textual data.[20] All this is done without the aid of predefined, explicit dictionaries or interpretive rules.

Using statistical associations of words in the patent abstract, 15 latent topics are identified as clusters of co-occurring words representing higher-order concepts. After a close look at the words associated with 15 topics (see Table 1), we identify a topic (Topic 14) that represents the use of technical artifacts because AI and related keywords will be part of technical artifacts. A detailed description of the composition of the words in Topic 14 is shown via word cloud in Fig. 2. We aim to see how AI usage in the patents that used technical artifacts changed during COVID-19. Thereby, all the patents that have Topic 14 with the highest probability are considered as a treated group of patents, and the rest become part of control group of patents.

We use an AI dictionary generated using Patentscope Artificial Intelligence Index by WIPO (shown in Table A1). This dictionary comprises keywords that indicate the presence of artificial intelligence in general. Each patent abstract is searched for the presence of words from the dictionary. To apply empirical text analytics to the patent abstracts, we create two variables, AI Presence and Count of AI Keywords. AI Presence is a binary variable coded as one if the patent abstract contains an AI-related keyword as matched from the AI dictionary, 0 otherwise. The Count of AI Keywords measures the number of AI-related keywords in the patent's abstract.

Table 1
Keywords corresponding to all topics obtained after LDA topic modeling.

Topics	Keywords
Topic 1	Antibodies, Protein, Polypeptide
Topic 2	Comprise, Material, Method, Particles
Topic 3	Cancer, Cell, Tumor
Topic 4	Methods, Treatment, Invention
Topic 5	Light, Image, Optical
Topic 6	Nucleic, Gene, DNA, RNA
Topic 7	Present, Composition, Invention, Formulation
Topic 8	Drug, Present, Improved
Topic 9	Sample, Method, Biological, Detection
Topic 10	Fluid, Liquid, Portion
Topic 11	Pharmaceutical, Composition, Invention, Therapeutic
Topic 12	Needle, Syringe, Device,
Topic 13	Invention, Virus, Vaccine, Immune, Infection
Topic 14	Device, Sensor, Data, System, Information, Signal
Topic 15	Method, Composition, Formulation, Using

Table A2 shows three different abstracts which we use to explain our methodology. In the first patent abstract, there is no presence of technical artifacts which we identify from Topic modeling. So, the first patent shall be part of control patents group. The second patent comprises words like computer, system, methods etc. which are part of technical artifacts as seen in Topic 14. This makes the second patent a treated patent but since there are no AI-related keywords, the value of AI Presence and Count of AI Keywords is 0. The third patent comprises Topic 14 words making it a treated patent, and also has AI-related keywords like machine learning, training dataset etc. We match these keywords from AI Dictionary and obtain values for AI Presence and Count of AI Keywords. Thus, within the treated patents (like second and third) we check how AI Presence and Count of AI Keywords change post COVID-19 compared to control patents.

As AI Presence is a limited dependent variable with binary values, we use a logistic model for its estimation along with the standard OLS. The Count of AI Keywords, on the other hand, has values ranging from 0 to 9. Because of high dispersion we also check the estimations with logarithm of Count of AI Keywords. To find the impact of AI on the pharmaceutical innovation process, we also generate a variable where we find number of days lapsed between application of patent and publication of the patent (see Table 2 for detailed variable description).

In Table 3, we show summary statistics of these variables by dividing them into two groups (treated and control group patents) and then specifying their statistics before and after January 2020. We assume that the impact of COVID-19 on pharmaceutical research started in January 2020 in US.[21] We use the difference-in-differences method as our empirical identification strategy. Fig. 3 shows the timeline diagram of our methodology. The empirical specification is shown below:

$$\begin{aligned} y_p = {} & \beta_0 + \beta_1 TreatedGroup_p + \beta_2 Covid_t + \beta_3 TreatedPatents_p \times Covid_t \\ & + Control + \delta_t + \theta_p + \gamma_f + \zeta_{fp} + \varepsilon \end{aligned} \tag{1}$$

The subscript p represents individual patents. Dependent variable y_p is used as AI Presence, Count of AI Keywords, Days App. to Pub. in three different estimations. $TreatedPatents_p$ corresponds to whether the patent comprises of technical artifacts group, identified by the presence of keywords in Topic 14 at the highest probability (1 for the treated group) or a nontechnical artifacts group (0 for the control group). $Covid_t$ would equal one if the patent were filed in months after which COVID-19 affected US (January 2020 to August 2020), zero otherwise.

Our coefficient of interest β_3 measures the change in the likelihood of the presence of AI keywords during COVID-19 when the dependent variable is AI Presence. When the dependent variable is Count of AI Keywords, β_3 measures change in the number of times dictionary

[18] We also replicate the results based on this sample to an extended period from January 2017 to December 2022. See Section 5.4 for details.

[19] We agree that there could be a generalizability concern since we use only US pharmaceutical industry data. However, note that US pharmaceutical market is the most innovative in the world, perhaps also the most vibrant. Though it may also be the most unregulated market, US patents are well used in prior work to draw generalizable trends. Prior work attest to the criticality of US patent data. For example, Singh et al. (2021) predict technology improvement rates for all technologies based on US patent data while Wagner and Wakeman (2016) verify patent-based measures with product commercialization using US-based pharmaceutical patents.

[20] LDAGIBBS command in STATA15 is used to obtain clustered keywords along with word probability.

[21] https://www.history.com/this-day-in-history/first-confirmed-case-of-coronavirus-found-in-us-washingtonstate.

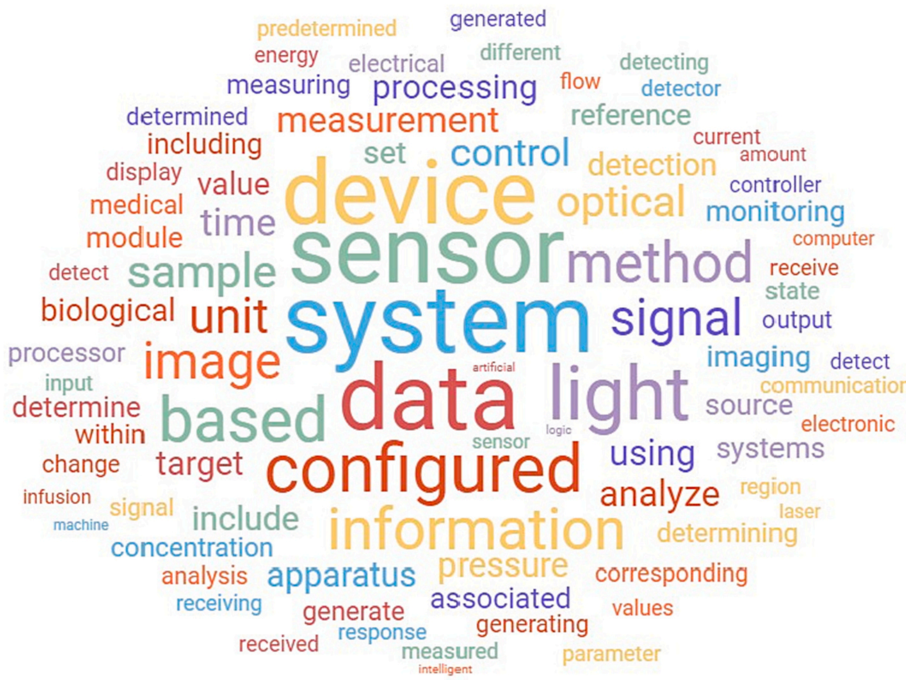

Fig. 2. Word cloud for topic 14. The figure represents all the major words that Topic 14 comprises. The larger the size of the word higher is the prominence of the keyword in the topic. We can see that the majority of the words highlight the technical artifact.

Table 2
Variable description.

Dependent variables	Definition and construction
AI presence	Coded as 1 if the patent abstract contains a AI related keyword, 0 if not present
Count of AI keywords	Number of AI related keywords present in the abstract of the patent
Log count of AI keywords	Logarithm of Count of AI Keywords
Log days App. to Pub.	Logarithm of number of days lapsed from application of the patent to publication of the patent
Independent variables	Definition and Construction
Treatment patent group	Coded as 1 if the patent abstract contains topic 14 obtained after LDA in highest probability, 0 otherwise
Covid	Coded as 1 if the patent application date is from January 2020, 0 otherwise
Control variable	Definition and Construction
Log number of words	Logarithm of number of words in the abstract of the patent

keywords are used in the patent abstract in treatment group as compared to the control group during pandemic. Positive and significant β_3 would indicate an increase in usage of AI in pharmaceutical research during pandemic. Similarly, when dependent variable is Days App. to Pub. β_3 measures change in number of days taken from application to publication of a patent. Negative and significant β_3 would indicate the reduced time taken for a pharmaceutical patent comprising AI to publish during pandemic.

In the specification, we control for the number of words in the abstract since there is a possibility of larger abstracts showing a higher count of AI keywords. We apply month dummies (δ_t) to account for time-varying common shocks, pharma subclass dummies (θ_p) are added to cover variations across different pharma subclasses, firm dummies (γ_f) are included to control for unobserved time-invariant heterogeneity across firms. We also add firm - pharma subclass dummies ζ_{fp} to account for unobserved heterogeneity across pharma subclasses within a firm. Robust standard errors clustered at topic level are used in all specifications.

The causal relation estimated from the difference-in-differences framework is based on the assumption of a parallel trend. It implies

Table 3
Summary statistics.

Treated patents				Control patents			
Oct 2019 to Dec 2020	N	Mean	Std. Dev.	Oct 2019 to Dec 2020	N	Mean	Std. Dev.
AI Presence	614	0.052	0.008	AI Presence	11,073	0.009	0.000
Count of AI Keywords	614	0.091	0.019	Count of AI Keywords	11,073	0.013	0.001
Log Count of AI Keywords	614	0.048	0.009	Log Count of AI Keywords	11,073	0.007	0.000
Log Days App. to Pub.	614	5.27	0.016	Log Days App. to Pub.	11,073	5.249	0.003
Log Number of Words	614	4.539	0.022	Log Number of Words	11,073	4.165	0.006
Jan 2020 to Aug 2020	N	Mean	Std. Dev.	Jan 2020 to Aug 2020	N	Mean	Std. Dev.
AI Presence	1227	0.064	0.007	AI Presence	17,763	0.009	0.000
Count of AI Keywords	1227	0.122	0.018	Count of AI Keywords	17,763	0.013	0.001
Log Count of AI Keywords	1227	0.062	0.007	Log Count of AI Keywords	17,763	0.007	0.000
Log Days App. to Pub.	1227	5.059	0.009	Log Days App. to Pub.	17,763	5.099	0.002
Log Number of Words	1227	4.521	0.015	Log Number of Words	17,763	4.160	0.004

Notes: The table shows the number of observations (N), mean and standard deviation separately for Treated and Control patents from October 2019 to August 2020 – with a split at January 2020 indicating pre- and post-treatment periods.

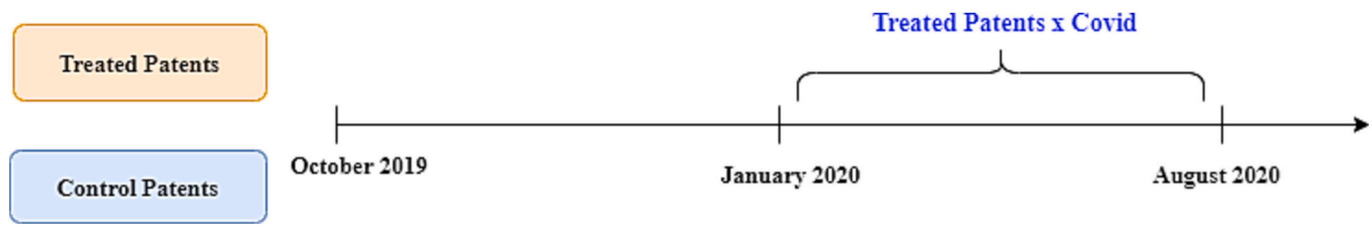

Fig. 3. Timeline diagram. The patents are classified as treated patents if the abstract comprises the topic (Topic 14) possessing technical artifacts, and the control group is all other patents. The study period is from October 2019 to August 2020. 'Covid' indicates months from January 2020 onwards.

that changes in outcome for the treatment group without the treatment would have been similar to the changes in the outcome in the control group. In other words, in the absence of COVID-19 shock, the presence and count of AI keywords should follow parallel trajectories over time. To validate this assumption for our study, we estimate the coefficient by plotting the pre-trends by using the event study design specification shown below:

$$y_p = \beta_0 + \sum_{t=Nov'19}^{Aug'20} \beta_t TreatedGroup_p + Control + \delta_t + \theta_p + \gamma_f + \zeta_{fp} + \varepsilon \tag{2}$$

where t ranges from November 2019 to August 2020, with October 2019 as the base year.

We plot the coefficients from the above equation for the outcome variables- AI Presence and Count of AI Keywords. Insignificant coefficients in the pre-period (till January 2020) would satisfy the assumption of parallel trends between the treatment and control group in our estimated results.

3.2. AIPD database

We crosscheck our findings using AIPD database from USPTO. This externally provided database by Giczy et al. (2022) helps us validate our conjecture from another data source, and also component-level analysis of AI enhances our inspection. The AIPD dataset is generated using machine learning models that use patent text, claims, and citations to identify eight AI component technologies - Machine Learning Model, Natural Language Processing, Evolutionary Computation Model, Knowledge Processing Model, AI Hardware Model, Planning/Control Model, Speech Model, and Vision Model. Through this algorithm, each patent receives a probability score between 0.0 and 1.0 indicating the presence of a technology component.[22] We select all the pharmaceutical patents in this database using CPC codes[23] from January 1980 to December 2020 and check how likelihood of presence of AI and all its components change during COVID-19 using below-mentioned specification:

$$y_p = \beta_0 + \beta_1 Covid_t + Control + \delta_t + \theta_p + \varepsilon \tag{3}$$

The subscript p represents individual patents. Dependent variable y_p is used as the likelihood of the presence of AI and its components and Log Days App. to Pub. in separate estimates. $Covid_t$ would equal one if the patent was filed in year 2020 after which COVID-19 affected US, zero otherwise. In the specification we control for number of words in the abstract since there is a possibility of larger abstracts showing higher count of AI keywords. Our coefficient of interest β_1 measures the change in the presence of AI and its components and time taken from application to publication of patents during pandemic in separate estimates. In the specification we apply year dummies (δ_t) to account for time-varying common shocks, pharma subclass dummies (θ_p) are added to cover variations across different pharma subclasses.

3.3. Leader vs. laggard firms - AI capability building

We further analyze the patent data to understand the mechanism through which an exogenous shock could propagate AI-related research. We first identify firms of all the patents in our baseline sample from October 2019 to August 2020. Within this set of firms, we mark the firms which have applied the patents where AI is present (as obtained from matching with the AI dictionary) and name these firms as leader firms and others as laggard firms. Next, we search for all the patents filed by the leader and laggard firms in the past two years, from September 2017 to September 2019. We look for the presence and count of AI keywords in the patents using the AI Dictionary in this time frame. We check if patents filed by the leader firms had better presence and higher count of AI keywords than the laggard firms. We use the below mentioned specification.

$$y_p = \beta_1 LeaderFirm_f + Control + \delta_t + \theta_p + \varepsilon \tag{4}$$

The subscript *p* represents individual patents. Dependent variable y_p is used as AI Presence and count of AI keywords in separate estimations. $LeaderFirm_f$ would equal one if the patent was filed by the leader firm, zero otherwise. Our coefficient of interest β_1 measures the presence and count of AI keywords used by leader firms as compared to laggard firms in last two years. In the specification we control for number of words in the abstract since there is a possibility of larger abstracts showing higher count of AI keywords. We apply month dummies (δ_t) to account for time-varying common shocks, pharma subclass dummies (θ_p) are added to cover variations across different pharma subclasses.[24] Robust standard errors clustered at firm level are used in all specifications.

4. Findings

4.1. Descriptive analysis

We show a simple trend of the average monthly presence of AI dictionary keywords in patent abstracts in treated and control group patents in Fig. 4. The left panel shows the average AI Presence, and the right panel indicates average Log Count of AI Keywords. The vertical line on January 2020 depicts the cut-off indicating the time after which COVID-19 effects were felt in the US. The smoothened lines in pre-treatment period from October 2019 to August 2020 in both panels are mostly parallel till January 2020 post when the dashed line in blue (representing treated group patents) starts rising. We can also see that the line in red representing control group remains mostly flat throughout the sample period. The sharp increase in the presence and count of AI dictionary keywords in patent abstracts of treated patents after the onset of COVID-19 highlights the increase in usage of innovative technologies in

[22] AI probability scores are translated into binary variables taking value one if the score $\geq$ 0.50 and zero otherwise.

[23] CPC: Cooperative Patent Classification. https://www.uspto.gov/web/patents/classification/cpc/html/cpc.html.

[24] Unlike previous specifications we don't use firm fixed effects since we wish to analyze firm-level changes and applying firm fixed effects would drop the leader firms variable.

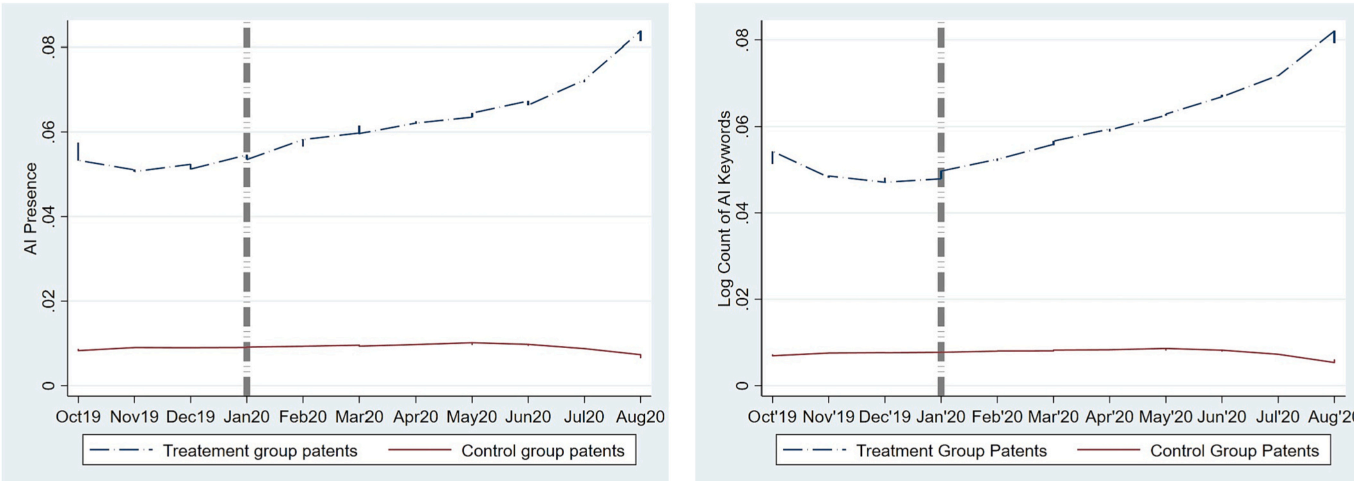

Fig. 4. Trend lines: figure on the left panel shows the monthly trends of the average AI Presence for treatment and control group patents. The figure on the right panel shows the monthly trends of the average Log Count of AI Keywords for treatment and control group patents. Treated patents in both figures are represented in blue dashed lines, while control patents are represented in solid red. The vertical grey lines in both figures indicate the commencement of COVID-19 in US in both panels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

pharmaceutical research in the US.

Table 3 corroborates the above descriptive findings with summary statistics pre and during COVID-19. For the control group, the mean value of AI presence remains unchanged at 0.009; on the other hand, for the treated patents group, it increases from 0.052 to 0.064 (23.07 % increase). Similarly, the Count of AI Keywords increases from 0.091 to 0.122 (34.06 % increase) for the treatment group, while for control group there is no change. Thus, descriptive analysis depicted via trend-line and summary statistics displays higher usage of AI techniques in pharmaceutical patents during the pandemic. We conduct a more systematic empirical analysis in the next sections to substantiate our research questions.

4.2. Main findings: Derwent database

In this section, we empirically evaluate the impact of COVID-19 on the usage of AI in pharmaceutical patents. Estimation of research question 1 using Eq. (1) is shown in Table 4, where the dependent variable is AI Presence in columns (1)–(3), Count of AI Keywords in columns (4)–(5) and the logarithm of Count of AI Keywords in column (6). Estimations in column (1) use logistic regression, and all other columns use ordinary least squares (OLS). Columns (1), (2) and (4) show the baseline estimations without any controls. In columns (3), (5) and (6) we introduce control for number of words in abstract, month dummies, pharma subclass dummies, firm dummies and firm-pharma subclass dummies.

Results in columns (1)–(3) show a positive and statistically significant increase in the likelihood of AI presence in the treated group after COVID-19. Causal estimates from logit estimation in column (1) reveal that the odds of AI's presence in pharmaceutical patents increased by 1.294 (exp (0.258)) after COVID-19. Results in column (2) indicate that the likelihood of the presence of AI increases by 1.3 % points in treated patents compared to control patents during pandemic. Same estimate goes up to 2.0 % points when all the controls are applied.

Columns (4)–(6) focus on the count of AI keywords. Results in column (4) indicate an increase of 0.031 in the count of AI keywords. With all controls, this goes up to 0.054 in column (5) which is 200 % increase over mean count of AI keywords (0.018) in our sample. In column (6)

Table 4
Change in presence and count of AI keywords in pharmaceutical patents post COVID-19.

	(1)	(2)	(3)	(4)	(5)	(6)
		AI presence			Count of AI keywords	
	LOGIT	OLS	OLS	OLS	OLS	Log count of AI keywords
Treated patents x Covid	0.258***	0.013***	0.020***	0.031***	0.054***	0.024***
	[0.073]	[0.001]	[0.003]	[0.001]	[0.007]	[0.003]
Treated patents	1.758***	0.043***	0.031***	0.078***	0.047***	0.027***
	[0.198]	[0.002]	[0.004]	[0.003]	[0.008]	[0.004]
Covid	−0.023	−0.000		0.000		
	[0.073]	[0.001]		[0.001]		
Log number of words			0.009***		0.018***	0.009***
			[0.002]		[0.005]	[0.002]
Month dummies	No	No	Yes	No	Yes	Yes
Pharma sub-class dummies	No	No	Yes	No	Yes	Yes
Firm dummies	No	No	Yes	No	Yes	Yes
Firm-pharma sub class dummies	No	No	Yes	No	Yes	Yes
Observations	30,527	30,527	30,527	30,527	30,527	30,527

Notes: The dependent variable in columns (1) to (3) is the likelihood of presence of AI while in columns (4) and (5) is Count of AI Keywords and in column (6) is Log Count of AI Keywords. In column (1) we use logistic regression estimated by maximum likelihood method. From column (2) to (6), we use ordinary least square method. Across model specifications, we see that the interaction term is positive and statistically significant. Thus, presence and intensity of AI increased significantly post COVID-19 for patents with technical artifacts. The time horizon is October 2019 to August 2020. The constant term is included but not reported. Robust standard errors clustered at topic level are presented in the parenthesis. '***','**','*' indicate significance at the 1 %, 5 % and 10 % respectively.

with logarithm of count of AI keywords as the dependent variable and with all controls in the specification, there is a significant increase of 2.4 % in count of AI keywords. Overall, empirical analysis on patents using AI Dictionary helps us claim that there has been an increase in usage of AI not just in presence but also in count.

We also checked for pre-trends in the data to establish the identification strategy. In Fig. 5, we present the coefficient estimates pre and during COVID-19. In the left panel, we generate a coefficient plot using Eq. (2) for AI Presence, and in the right panel, we do the same with Log Count of AI Keywords as the dependent variable. The horizontal red line indicates zero coefficient, indicating no significant difference between the treated and control groups. The vertical red line indicates initiation of treatment in January 2020. We can see that before January the trend line in both figures was close to zero which indicates the absence of any significant pre-trends between treated and control groups. We find a sharp change in coefficient estimates as the pandemic began in January 2020.

4.3. Increasing speed of publication

We have established the increase in AI usage in pharmaceutical patents with the onset of COVID-19. In this section, we seek to evaluate the impact of this increase on the speed of publication, especially given long time frames involved in the pharmaceutical innovation process. In Fig. A1 we show how the time taken from application to publication of all pharmaceutical patents changes month-on-month from 2017 to 2022. We see that post COVID-19 there is a significant dip in the time taken from application to publication for patents with AI.

Table 5 empirically evaluates research question 2 and presents the results of the regression analyses by estimating Eq. (1). The dependent variable in all columns is the logarithm of the number of days from the application date to publication date. In all columns we apply ordinary least square method. Column (1) shows the baseline result without any fixed effects. In column (2) we introduce control for number of words in abstract, month dummies and pharma subclass dummies. In column (3) we add firm dummies and firm-pharma subclass dummies.

Results in columns (1)–(3) show a negative and statistically significant decrease in the days it takes from application to publication of patent during COVID-19. Causal estimates from baseline estimation in column (1) reveal that number of days decreased by 6.0 % after COVID-19. With all the controls in column (3) same estimate is 5.4 %.[25] Overall, we find the presence of AI is nudging the authorities to fast-track the publication process. One possible reason is that it may be because they find more value in AI-driven patents in the future. A 5 % decrease in the overall innovation process in a world where new drugs take 12 years to hit the market seems non-trivial and perhaps will only improve in the coming days.

4.4. Crosschecking using AIPD database

Table 6 presents the results of the regression analyses by estimating Eq. (3). In column (1), the dependent variable is the overall AI Presence. In column (2), the dependent variable is Log Days App. to Pub. and dependent variables from column (3) to column (9) are different components of AI. In all columns, we apply ordinary least square method with year and pharma subclass dummies.

In column (1), we find a significant increase in the presence of AI in pharmaceutical patents during the pandemic. In column (2), we find a significant decrease in time taken from application to publication in pharmaceutical patents during COVID-19. Similarly, all AI components show a significant increase during COVID-19 in the range of 0.4 % points to 3.6 % points except for speech model and natural language processing. Thus, these results help us triangulate our main findings using a different dataset to strengthen our claim that there was an increase in AI in pharmaceutical patents during COVID-19.

4.5. Leader firms build capability in advance

Table 7 presents the results of the regression analyses for research question 3 by estimating Eq. (4). The dependent variable is AI Presence in columns (1) and (2), the Count of AI Keywords in column (3), and the logarithm of Count of AI Keywords in column (4). Estimations in column (1) use logistic regression, and all other columns use ordinary least squares (OLS). Column (1) shows the baseline estimations controlling for the number of words in the abstract without any fixed effects. In column (2) - column (4) we introduce month dummies and pharma subclass dummies.

Results in Table 7 estimate the difference in the presence and count of AI keywords in patents filed by leader firms compared to laggard firms two years before the pandemic. We would refer to columns (2) and column (4) for interpretation as these models are most conservative with full controls. Column (2) where we evaluate AI Presence, shows a positive and significant coefficient ($\beta = 0.077$), indicating that likelihood of presence of AI in patents filed by leader firms was 7.7 % points higher than the laggard firms. Similarly, findings in column (4), where we look at count of AI keywords, indicate a positive and significant coefficient ($\beta = 0.102$). Thus, count of AI keywords was higher by 10.2 % in pharmaceutical patents of leader firms compared to laggard firms. Therefore, this analysis indicates that the firms that adopted AI quickly during the pandemic could do so because of the AI capability they had built in previous years, also suggesting an absorptive capacity argument to the AI adoption process in biopharmaceutical innovation.

5. Robustness checks

5.1. Coarsened exact matching (CEM)

CEM acts as a means for robustness test for differences in differences estimates where researcher arbitrariness in choosing control groups could create biased results (Iacus et al., 2012). CEM improves causal estimation by narrowing down the imbalance in covariates among treated and control groups. "In coarsened exact matching, we temporarily coarsen the data, exact match on these coarsened data, and then run their analysis on the uncoarsened, matched data" (Blackwell et al., 2009). Coarsened exact matching is fast, easy to use, requires fewer assumptions, is easily automated, and possesses better statistical properties than existing matching methods. These advantages of the coarsened exact matching method have led to the use of this method in recent studies (see Chen et al. (2022); Fry (2021); Wang and Zheng (2022)).

In this study, our interest outcomes are AI Presence, Log count of AI keywords, and Log Days App. to Pub. We match treatment and control groups of patents based on the number of words in the patent abstract, pharma subclass, and cited patents. Estimation using the CEM approach is shown in Table 8. We can see a slight drop in the number of observations in the CEM approach as it generates causal analysis on uncoarsened control group. We can see that the interaction coefficient obtained from CEM has same or improved magnitude and significance as compared with our main estimates. CEM lends strength to our identified causal impact of COVID-19 on the increase in usage of AI in pharmaceutical patents.

5.2. Falsification test with 2018–19 sample

As shown in Fig. 3, our baseline results timeline is from October 2019 to August 2020, with the COVID-19 shock coming in January 2020. In this exercise, we replicate all our estimations with a changed timeline

[25] We also generate coefficient plot using Eq. (2) with Log Days from App. to Pub. as the dependent variable and find a sharp decrease in interaction coefficient estimates as the pandemic began in January 2020. The plot can be made available on request.

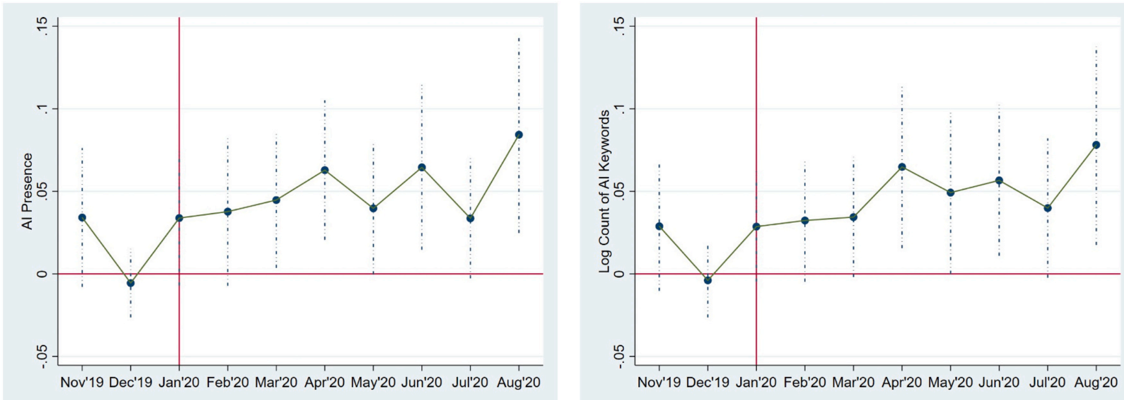

Fig. 5. Event study design: left panel shows the fully specified coefficient plot for AI Presence, showing difference-in-differences coefficients comparing the treatment and control group of patents for each month. The right panel shows the difference-in-differences coefficients for each month for the Log Count of AI Keywords estimation. The vertical red line indicates the commencement of COVID-19 in the US in both panels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Reduction in time from application to publication of patents post COVID-19.

DV: log days App. to Pub.	(1)	(2)	(3)
Treated patents x Covid	−0.060***	−0.055***	−0.054**
	[0.019]	[0.019]	[0.025]
Treated patents	0.020	0.032*	0.052**
	[0.017]	[0.017]	[0.022]
Covid	−0.151***		
	[0.005]		
Log number of words		−0.006*	0.001
		[0.003]	[0.005]
Month dummies	No	Yes	Yes
Pharma sub-class dummies	No	Yes	Yes
Firm dummies	No	No	Yes
Firm-pharma sub class dummies	No	No	Yes
Observations	30,527	30,527	30,527

Notes: The dependent variable in all columns is logarithm of days taken from application of patent to publication. Negative and statistically significant interaction coefficient indicates that time taken from application to publication significantly reduced post COVID-19 in treated patents. The time horizon is October 2019 to August 2020. The constant term is included but not reported. Robust standard errors are presented in the parenthesis. '***','**','*' indicate significance at the 1 %, 5 % and 10 % respectively.

exactly a year before, from October 2018 to August 2019, with an alternate placebo treatment from January 2019. Estimations are shown in Table 9. We find that all interaction coefficients in all estimations are insignificant. Thus, keeping a similar time trend month-on-month, a placebo treatment one year before doesn't cause changes. This analysis helps us establish the exogenous nature of COVID-19 shock and the effects caused by it. This analysis also helps us to rule out existing pre-trends (if any).

5.3. *Alternate control group*

In all our analyses till now, we took treated patents as the ones that had technical artifacts identified by Topic 14, and all other patents became part of the control group. In this section we revise our control group to be comprised of the topics that are closer to Topic 14. Thus, the new control group comprises patents with topics 2, 4, 7 and 15 in highest probability. Results of this analysis using Eq. (1) is shown in Table 10. We can see that all interaction coefficients remain positive and significant even with a stringent sub-sample analysis.

5.4. *Extended time period*

In this section, we replicate our baseline results on the extended data from January 2017 to December 2022. The purpose of choosing this sample was to select an extended period of three years before and three years during COVID-19. Table 11 presents the results of the regression analyses for extended time by estimating Eq. (1). We find that all our baseline results hold in the long term. These results act as a robustness check to allay the concern about the time-enduring nature of our baseline results.

Table 6
Crosschecking with AIPD database the presence of AI components in pharmaceutical patents post COVID-19.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Overall AI presence	Log days App. to Pub.	Knowledge processing model	Planning/control model	Vision model	Machine learning model	Evolutionary computation model	AI hardware model	Speech model	Natural language processing
Covid	0.048***	−1.317***	0.036***	0.025***	0.017***	0.008**	0.004**	0.004*	0.000	−0.000
	[0.010]	[0.016]	[0.007]	[0.006]	[0.006]	[0.004]	[0.002]	[0.002]	[0.000]	[0.000]
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pharma sub-class dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	482,075	482,075	482,075	482,075	482,075	482,075	482,075	482,075	482,075	482,075

Notes: The dependent variable in columns (1) is the likelihood of overall presence of AI in pharma patents. In column (2) it is the logarithm of days taken from application of patent to publication. Dependent variable in column (3)–(9) are different components of AI. We find that in all components of AI except for speech and NLP there is a positive and significant change post COVID-19. The time horizon is January 1980 to December 2020. The constant term is included but not reported. Robust standard errors are presented in the parenthesis '***','**','*' indicate significance at the 1 %, 5 % and 10 % respectively.

Table 7
Tracing capability building between leader and laggard firms.

	(1)	(2)	(3)	(4)
	AI presence LOGIT	AI presence OLS	Count of AI keywords	Log count of AI keywords
Leader firm	0.318***	0.077***	0.316***	0.102***
	[0.018]	[0.004]	[0.028]	[0.006]
Log number of words	0.637***	0.129***	0.853***	0.214***
	[0.011]	[0.002]	[0.024]	[0.003]
Month dummies	No	Yes	Yes	Yes
Pharma subclass dummies	No	Yes	Yes	Yes
Observations	81,958	81,958	81,958	81,958

Notes: The dependent variable in columns (1) and (2) are AI Presence and in columns (3) and (4) are Count of AI Keywords and Logarithm of Count of AI Keywords respectively. In column (1) we use logistic regression estimated by maximum likelihood method. From column (2) to (4), we use ordinary least square method. Across model specifications, we see that the coefficient of leader firm is positive and statistically significant. Thus, firms that were adopting AI quickly post COVID-19 could do so because they had already built-up capability in past years. The time horizon is September 2017 to September 2019. The constant term is included but not reported. Robust standard errors clustered at firm level are presented in the parenthesis. '***','**','*' indicate significance at the 1 %, 5 % and 10 % respectively.

Table 8
Robustness check: coarsened exact matching.

	(1)	(2)	(3)
	AI presence	Log count of AI keywords	Log days App. to Pub.
Treated patents x Covid	0.020***	0.024***	−0.069***
	[0.004]	[0.003]	[0.010]
Treated patents	0.030***	0.026***	0.056***
	[0.004]	[0.004]	[0.012]
Log number of words	0.010***	0.012***	0.008
	[0.002]	[0.003]	[0.005]
Month dummies	Yes	Yes	Yes
Pharma sub-class dummies	Yes	Yes	Yes
Firm dummies	Yes	Yes	Yes
Firm-pharma sub class dummies	Yes	Yes	Yes
Observations	29,924	29,924	29,924

Notes: The dependent variable in columns (1), (2) and (3) are AI Presence, Count of AI Keywords and Log Days App. to Pub. We use ordinary least square method in all models. We can see that the interaction coefficient for AI Presence and Count of AI Keywords are positive and significant and that of Log Days App. to Pub. is negative and significant as in the baseline results. Thus, when we match the treatment and control groups based on pharma class, number of words in abstract and citation of patents we see that all the results improve. The time horizon is October 2019 to August 2020. The constant term is included but not reported. Robust standard errors clustered at topic level are presented in the parenthesis. '***','**','*' indicate significance at the 1 %, 5 % and 10 % respectively.

Table 9
Falsification check: sample of 2018–19.

	(1)	(2)	(3)
	AI presence	Log count of AI keywords	Log days App. to Pub.
Treated patents x Covid	0.009	0.056	−0.032
	[0.011]	[0.056]	[0.029]
Treated patents	−0.015	−0.023	0.059**
	[0.022]	[0.050]	[0.025]
Log number of words	0.087***	0.685***	0.016***
	[0.015]	[0.012]	[0.006]
Month dummies	Yes	Yes	Yes
Pharma sub-class dummies	Yes	Yes	Yes
Firm dummies	Yes	Yes	Yes
Firm-pharma sub class dummies	Yes	Yes	Yes
Observations	24,643	24,643	24,643

Notes: The dependent variable in columns (1), (2) and (3) are AI Presence, Count of AI Keywords and Log Days App. to Pub. We use ordinary least square method in all models. Across model specifications, we see that the interaction coefficient is insignificant. Thus, these results provide a placebo check and show that results don't hold exactly one year back. The time horizon is October 2018 to August 2019 with placebo COVID dummy starting from January 2019. The constant term is included but not reported. Robust standard errors are presented in the parenthesis. '***','**','*' indicate significance at the 1 %, 5 % and 10 % respectively.

Table 10
Robustness check: alternate control group.

	(1)	(2)	(3)
	AI presence	Log count of AI keywords	Log days App. to Pub.
Treated patents x Covid	0.019*	0.023**	−0.041**
	[0.007]	[0.006]	[0.015]
Treated patents	0.034**	0.028**	0.043
	[0.008]	[0.006]	[0.027]
Log number of words	0.007	0.007	0.004
	[0.004]	[0.004]	[0.012]
Month dummies	Yes	Yes	Yes
Pharma sub-class dummies	Yes	Yes	Yes
Firm dummies	Yes	Yes	Yes
Firm-pharma sub class dummies	Yes	Yes	Yes
Observations	11,923	11,923	11,923

Notes: The dependent variable in columns (1), (2) and (3) are AI Presence, Count of AI Keywords and Log Days App. to Pub. We use ordinary least square method in all models. We can see that the interaction coefficient for AI Presence and Count of AI Keywords are positive and significant and that of Log Days App. to Pub. is negative and significant as in the baseline results. Thus, when we restrict the sample to keep control groups that comprise of topics that closely match with topic in treatment group we find all results hold. The time horizon is October 2019 to August 2020. The constant term is included but not reported. Robust standard errors clustered at topic level are presented in the parenthesis. '***','**','*' indicate significance at the 1 %, 5 % and 10 % respectively.

6. Discussion

Research and development in pharmaceutical firms recently are going through productivity crises (Cockburn, 2006; Pammolli et al., 2011). Importantly, there is additional pressure on the pharmaceutical sector to produce drugs to curb the spread of evolving new variants of COVID-19. At this point, the role of emerging technologies such as data analytics and artificial intelligence is undeniable for 'ultrafast' innovation. AI is seen as an important tool providing a long-term and sustainable solution (Aristodemou, 2020; Yamashita et al., 2021). AI has a role in targeting drug discovery (Zielinski, 2021), in designing modality (Colombo, 2020), in designing preclinical experiments (Shanbhogue et al., 2021) and in choosing patients for clinical trials (Smalley, 2017). Thus AI could enter into drug development life cycle at every step.

Though the benefits of AI are long known, its uptake by pharmaceutical firms has been appalling. In 2019, KPMG surveyed Fortune 500 companies to find only 17 % of firms using AI or ML at scale (KPMG, 2019).[26] We study if the situation changed during COVID-19 and if there is a change in the way innovation is done in pharmaceutical industry, especially given broader trends in digitalization of the average firm,

[26] "https://advisory.kpmg.us/content/dam/advisory/en/pdfs/2019/8-ai-trends-transforming-the-enterprise.pdf".

Table 11
Extended time period: January 2017–December 2022.

	(1)	(2)	(3)
	AI presence	Log count of AI keywords	Log days App. to Pub.
Treated patents x Covid	0.024***	0.047***	−0.052***
	[0.007]	[0.017]	[0.009]
Treated patents	0.073**	0.259***	0.056***
	[0.025]	[0.012]	[0.007]
Log number of words	0.105***	0.259***	0.002
	[0.012]	[0.004]	[0.002]
Month dummies	Yes	Yes	Yes
Pharma sub-class dummies	Yes	Yes	Yes
Firm dummies	Yes	Yes	Yes
Firm-pharma sub class dummies	Yes	Yes	Yes
Observations	141,379	141,379	141,173

Notes**:** The dependent variable in columns (1), (2) and (3) are AI Presence, Log Count of AI Keywords and Log Days App. to Pub. We use ordinary least square method in all models. Across model specifications, we see that the interaction coefficient is significant. Thus, these results provide a check and show that results hold even if we extend the time period. The time horizon is January 2017 to December 2022 with COVID dummy starting from January 2020. The constant term is included but not reported. Robust standard errors are presented in the parenthesis. '***','**','*' indicate significance at the 1 %, 5 % and 10 % respectively.

work from home models and technology enabled AI diffusion in manufacturing sector during COVID-19.

To investigate this, we use LDA based topic modeling technique on patent abstracts to identify patents with technical artifacts as treated group patents. Then using the AI dictionary, we observe AI usage change in these patents. Using difference-in-differences methodology we find a significant increase in AI Presence and Count of AI Keywords in treated group patents compared to control group patents. We also find that with the usage of AI in treated patents there is a significant decrease in time taken from application to publication of the patent. All these results are cross-checked using two different databases.

In further investigations, we search for the catalyst behind this sudden change in innovation priorities to find that these results are driven by firms that had already built AI capability in the past. Our research highlights that the state of AI capability for pharmaceutical firms and the pandemic acted as a digital accelerator for embracing AI methodologies for pharmaceutical research (Jacobides et al., 2021).

Contrary to general perception, there is a body of prior work that mentions that AI adoption in healthcare has lagged when compared to other industries (Goldfarb et al., 2020; Stempniak, 2022). AI can offer unprecedented opportunities, but there are barriers to AI adoption in healthcare, some recent scholars have argued (Goldfarb and Teodoridis, 2022). There is also a discussion on major hurdles in diffusion of AI in healthcare at managerial and regulatory levels (Agrawal et al., 2022). Interestingly, success of AI depends on the physicians whose thriving is threatened by AI (Dranove and Garthwaite, 2022). Our paper possibly provides the first evidence of how AI usage has increased in pharmaceutical patents and is in the right direction to bring change in healthcare contrary to what this prior work has argued.

Since we now find an increase in the usage of AI-based approaches in the pharmaceutical space, policymakers should effectively formulate policies that will govern the 'dark side' of AI (Cao et al., 2021) and ensure the ethics of AI in the pharmaceutical sector. Key potential concerns caused by AI are the trustworthiness of decisions by AI technologies, the privacy of data, and the cybercrime threat (Carter, 2020). Formal regulation of AI technologies is necessary to ensure that already established human rights and the legal and ethical principles of society are not contravened (Carter, 2020).

In healthcare, the concern about AI is more serious because AI systems learn to make decisions based on training data, which can include biased human decisions or reflect historical or social inequities (Manyika, 2019). A recent survey found that 75 % of healthcare insiders are worried that AI could threaten the privacy and security of patient data (Samantha, 2020). Another ethical concern with AI is people losing their jobs and whether the machines should replace healthcare workers, where feelings, empathy, and warmth are very important factors (Kostic et al., 2019).

At the firm level, while considering the ethical concerns is important, it shall be prudent to fund advanced technology research to stay future-ready. A recent study by Boutillier et al. (2023) substantiates our results as they explain how Sanofi, a leading French pharmaceutical company, lagged behind Pfizer, AstraZeneca, and Moderna in developing COVID-19 vaccine because of reluctance in usage of modern technology in the past. Thus, though COVID-19 nudged the pharmaceutical industry towards adoption of AI, all firms were not equally responsive. This is an important finding to highlight in the conversation on absorptive capacity around firm-level AI adoption that we are among the first to unearth.

Since we find that AI patents are getting approved faster and are percolating into pharmaceutical patents, it may mean that there is an increase in the R&D efficiencies in the pharmaceutical sector. An increase in R&D efficiencies with quicker patent approvals mean that firms can cut down on R&D costs and this could impact prices of medicines in a direction that is socially welfare-enhancing.

With all its merit, we agree that our baseline results are obtained at a partial equilibrium stage, and the data we observe is only based in the US. One can always include patents from European Union, China, Japan or other important global jurisdictions and markets for broader analysis. Future researchers can also build upon this analysis to observe in detail the benefits or drawbacks that increasing AI usage can provide. Much more, therefore, remains to be done.

CRediT authorship contribution statement

Sawan Rathi: Conceptualization, Data curation, Software, Formal analysis, Investigation, Writing - original draft, Writing - review & editing. **Adrija Majumdar:** Conceptualization, Validation, Formal analysis, Investigation, Software, Supervision, Writing - review & editing. **Chirantan Chatterjee:** Conceptualization, Validation, Formal analysis, Investigation, Funding acquisition, Project administration, Supervision, Visualization, Writing - review & editing.

Data availability

The authors do not have permission to share data.

Appendix A

Table A1
AI dictionary keywords.

3d imaging	image segmentation	phonological analysis	speech-to-speech
active vision	information extraction	phonology	speech-to-text
activity recognition	interest point detection	predictive	stemming
artificial intelligence	knowledge representation	prescriptive	synthesizer
augmented reality	lemmatization	recommender	text analytics
biometrics	machine translation	reconstruction	text mining
camera calibration	mechatronic	robotic	text-to-speech
chatbot	mobile agent	salient region detection	text-to-voice
computational control	morphological	scene anomaly detection	texture representations
computational photography	morphology	scene understanding	video segmentation
computer vision	motion capture	semantic	video summarization
content extraction	motion path	semantics	virtual reality
data-mining	multi-agent	shape inference	visual
distributed ai	natural language processing	shape representation	visual indexing
epipolar geometry	object detection	speech processing	visual inspection
hierarchical representation	object identification	speech recognition	visual retrieval
hyperspectral imaging	object recognition	speech synthesis	voice recognition
image representation	personal assistant	speech-generating	voice-to-text

Notes**:** Table shows excerpts of the keywords in the AI dictionary that we use to match in the patent abstract to identify if AI is present and to what extent.

Table A2
Patent abstract example.

Patent Number: US10322193B1; Application Date: 2019-06-18.
Abstract: The present invention relates to therapeutic conjugates with improved ability to target various diseased cells containing a targeting moiety (such as an antibody or antibody fragment), a linker and a therapeutic moiety, and further relates to processes for making and using the conjugates.
Treated Patent: No; AI Presence: No; Count of AI Keywords: 0.
Patent Number: US10395761B1; Application Date: 2019-04-22.
Abstract: Various methods, systems, computer readable media, and graphical user interfaces (GUIs) are presented and described that enable a subject, doctor, or user to characterize or classify various types of cancer precisely. Additionally, described herein are methods, systems, computer readable media, and GUIs that enable more effective specification of treatment and improved outcomes for patients with identified types of cancer. Some embodiments of the methods, systems, computer readable media, and GUIs described herein comprise obtaining RNA expression data and/or whole exome sequencing (WES) data for biological samples; determining a respective plurality of molecular-functional (MF) profiles for a plurality of subjects; clustering the plurality of MF profiles to obtain MF profile clusters; determining a molecular-functional (MF) profile for an additional subject; and identifying, from among the MF profile clusters, a particular MF profile cluster with which to associate the MF profile for the subject.
Treated Patent: Yes; AI Presence: No; Count of AI Keywords: 0.
Patent Number: US2020395100A1; Application Date: 2020-08-26.
Abstract: Systems and methods are disclosed for generating a therapeutic response predict or detecting a disease, by: using a genetic analyzer to generate genetic information; receiving into computer memory a training dataset comprising, for each of a plurality of individuals having a disease, (1) genetic information from the individual generated at first time point and (2) treatment response of the individual to one or more therapeutic interventions determined at a second, later, time point; and implementing a machine learning algorithm using the dataset to generate at least one computer implemented classification algorithm, wherein the classification algorithm, based on genetic information from a subject, predicts therapeutic response of the subject to a therapeutic intervention.
Treated Patent: Yes; AI Presence: Yes; Count of AI Keywords: 5

Notes: Table shows three examples of patent abstracts along with the patent number and application date. The first abstract doesn't belong to a treated patents as it doesn't contain any technical artifact in the abstract. The second abstract is the treated patent because it comprises words like computer, system, method etc. but doesn't contain any AI related words. Third abstract is from a treated patent and contains AI keywords. We also show values of AI Presence and Intensity that are obtained post this matching in all patents.

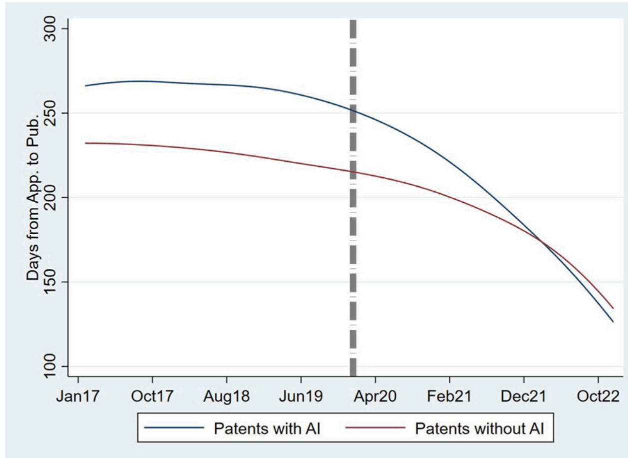

Fig. A1. Days from application to publication of patent.

References

Abramo, G., D'Angelo, C.A., Di Costa, F., Solazzi, M., 2011. The role of information asymmetry in the market for university–industry research collaboration. J. Technol. Transfer 36 (1), 84–100.

Acemoglu, D., Restrepo, P., 2020. The wrong kind of AI? Artificial intelligence and the future of labour demand. Camb. J. Reg. Econ. Soc. 13 (1), 25–35.

Agrawal, A., Gans, J.S., Goldfarb, A., 2019. Exploring the impact of artificial intelligence: prediction versus judgment. Inf. Econ. Policy 47, 1–6.

Agrawal, A., Gans, J.S., Goldfarb, A., et al., 2022. Similarities and Differences in the Adoption of General Purpose Technologies. National Bureau of Economic Research.

Amabile, T.M., 2020. Creativity, artificial intelligence, and a world of surprises. Acad. Manag. Discov. 6 (3), 351–354.

Aristodemou, L., 2020. Identifying Valuable Patents: A Deep Learning Approach (Unpublished doctoral dissertation). University of Cambridge.

Aristodemou, L., Tietze, F., 2018. The state-of-the-art on intellectual property analytics (IPA): A literature review on artificial intelligence, machine learning and deep learning methods for analysing intellectual property (IP) data. World Patent Inf. 55, 37–51.

Babu, K.A., Shirlin, M., Manjula Devi, A., 2021. Artificial intelligence in pharma. Int. J. Curr. Res. Physiol. Pharmacol. 5 (1), 1–3.

Bada, A., 2019. World's Oldest Pharmaceutical Merck Wins New AI & Blockchain Patent. BTCNN.

Bessant, J., Kunne, C., M˙oslein, K., 2012. Opening up Healthcare Innovation: Innovation Solutions for a 21st Century Healthcare System. Advanced Institute for Management Research.

Bianchini, S., Muller, M., Pelletier, P., 2022. Artificial intelligence in science: an emerging¨ general method of invention. Res. Policy 51 (10), 104604.

Blackwell, M., Iacus, S., King, G., Porro, G., 2009. CEM: coarsened exact matching in Stata. Stata J. 9 (4), 524–546.

Bloch, H., Metcalfe, S., 2018. Innovation, creative destruction, and price theory. Ind. Corp. Chang. 27 (1), 1–13.

Block, J., 2020. COVID-19 puts spotlight on artificial intelligence: pandemic amplifies the power of AI to generate actionable information. GEN Edge 2 (1), 149–155.

Boutillier, S., Laperche, B., Lebert, D., Elouaer-Mrizak, S., 2023. A systemic analysis of the technological trajectory at company level based on patent data: the case of Sanofi's vaccine technology. Technovation 124, 102746.

Bresnahan, T., 2010. General purpose technologies. In: Handbook of the Economics of Innovation, 2, pp. 761–791.

Bresnahan, T.F., Trajtenberg, M., 1995. General purpose technologies 'engines of growth'? J. Econ. 65 (1), 83–108.

Brynjolfsson, E., Rock, D., Syverson, C., 2021. The productivity J-curve: how intangibles complement general purpose technologies. Am. Econ. J. Macroecon. 13 (1), 333–372.

Bughin, J., Hazan, E., Lund, S., Dahlstrom, P., Wiesinger, A., Subramaniam, A., 2018. Skill shift: automation and the future of the workforce. McKinsey Glob. Inst. 1, 3–84.

Burki, T., 2019. Pharma blockchains AI for drug development. Lancet 393 (10189), 2382.

Calvino, F., Samek, L., Squicciarini, M., Morris, C., 2022. Identifying and Characterising AI Adopters: A Novel Approach Based on Big Data. OECD.

Cao, G., Duan, Y., Edwards, J.S., Dwivedi, Y.K., 2021. Understanding managers' attitudes and behavioral intentions towards using artificial intelligence for organizational decisionmaking. Technovation 106, 102312.

Cardinal, L.B., 2001. Technological innovation in the pharmaceutical industry: the use of organizational control in managing research and development. Organ. Sci. 12 (1), 19–36.

Carter, D., 2020. Regulation and ethics in artificial intelligence and machine learning technologies: where are we now? Who is responsible? Can the information professional play a role? Bus. Inf. Rev. 37 (2), 60–68.

Chan, H.S., Shan, H., Dahoun, T., Vogel, H., Yuan, S., 2019. Advancing drug discovery via artificial intelligence. Trends Pharmacol. Sci. 40 (8), 592–604.

Chen, W., Zaiyan, W., Xie, K., 2022. The battle for homes: how does home sharing disrupt local residential markets? Manag. Sci. 68 (12), 8515–9218.

Cockburn, I.M., 2006. Is the pharmaceutical industry in a productivity crisis? Innov. Policy Econ. 7, 1–32.

Cockburn, I.M., Kortum, S.S., Stern, S., 2002. Are all Patent Examiners Equal? The Impact of Examiner Characteristics. National Bureau of Economic Research.

Cockburn, I.M., Henderson, R., Stern, S., 2018. The impact of artificial intelligence on innovation: An exploratory analysis. In: The Economics of Artificial Intelligence: An Agenda. University of Chicago Press, pp. 115–146.

Colombo, S., 2020. Applications of artificial intelligence in drug delivery and pharmaceutical development. In: Artificial Intelligence in Healthcare. Elsevier, pp. 85–116.

Dal Mas, F., Biancuzzi, H., Massaro, M., Miceli, L., 2020. Adopting a Knowledge Translation Approach in Healthcare Co-production. A Case Study. Management Decision.

Davenport, T.H., Ronanki, R., 2018. Artificial intelligence for the real world. Harv. Bus. Rev. 96 (1), 108–116.

Dernis, H., Calvino, F., Moussiegt, L., Nawa, D., Samek, L., Squicciarini, M., 2023. Identifying Artificial Intelligence Actors Using Online Data. OECD.

DiMasi, J.A., Grabowski, H.G., Hansen, R.W., 2016. Innovation in the pharmaceutical industry: new estimates of R&D costs. J. Health Econ. 47, 20–33.

Dranove, D., Garthwaite, C., 2022. Artificial Intelligence, the Evolution of the Healthcare Value Chain, and the Future of the Physician. National Bureau of Economic Research.

Erzurumlu, S.S., Pachamanova, D., 2020. Topic modeling and technology forecasting for assessing the commercial viability of healthcare innovations. Technol. Forecast. Soc. Chang. 156, 120041.

Fry, C.V., 2021. Crisis and the trajectory of science: evidence from the 2014 Ebola outbreak. Rev. Econ. Stat. 1–35.

Giczy, A.V., Pairolero, N.A., Toole, A.A., 2022. Identifying artificial intelligence invention: a novel AI patent dataset. J. Technol. Transf. 47 (2), 476–505.

Goldfarb, A., Teodoridis, F., 2022. Why Is AI Adoption in Health Care Lagging. Brookings.

Goldfarb, A., Taska, B., Teodoridis, F., 2020. Artificial intelligence in health care? Evidence from online job postings. AEA Papers Proc. 110, 400–404.

Green, C., 2017. Technology innovation will help prepare drug discovery for smart screening era. Drug Discov. 18, 17.

Grimaldi, M., Quinto, I., Rippa, P., 2013. Enabling open innovation in small and medium enterprises: a dynamic capabilities approach. Knowl. Process. Manag. 20 (4), 199–210.

Haefner, N., Wincent, J., Parida, V., Gassmann, O., 2021. Artificial intelligence and innovation management: a review, framework, and research agenda. Technol. Forecast. Soc. Chang. 162, 120392.

Hammadou, H., Paty, S., Savona, M., 2014. Strategic interactions in public R&D across European countries: a spatial econometric analysis. Res. Policy 43 (7), 1217–1226.

Iacus, S.M., King, G., Porro, G., 2012. Causal inference without balance checking: coarsened exact matching. Polit. Anal. 1–24.

Igna, I., Venturini, F., 2023. The determinants of AI innovation across European firms. Res. Policy 52 (2), 104661.

Irving, W.B., 2020. What history tells us about the accelerating AI revolution. Wall Street J. https://www.wsj.com/articles/what-history-tellsus-about-the-accelerating-ai-revolution-01590169213?reflink=desktopwebshare_permalink.

Jacobides, M.G., Brusoni, S., Candelon, F., 2021. The evolutionary dynamics of the artificial intelligence ecosystem. Strat. Sci. 6 (4), 412–435.
Kim, Y.K., Oh, J.B., 2017. Examination workloads, grant decision bias and examination quality of patent office. Res. Policy 46 (5), 1005–1019.
Kostic, E.J., Pavlovi'c, D.A., Zivkovic, M.D., 2019. Applications of artificial intelligence′ in medicine and pharmacy: ethical aspects. Acta Med. Medianae 58 (3), 128–137.
KPMG, 2019. AI Transforming the Enterprise - 8 Key Adoption Trends.
Kramme, S., 2021. Navigating the new normal: which firms have adapted better to the COVID-19 disruption? Technovation 102368.
Kummar, S., Kinders, R., Rubinstein, L., Parchment, R.E., Murgo, A.J., Collins, J., 2007. Compressing drug development timelines in oncology using phase '0' trials. Nat. Rev. Cancer 7 (2), 131–139.
Lee, S.M., Trimi, S., 2021. Convergence innovation in the digital age and in the COVID-19 pandemic crisis. J. Bus. Res. 123, 14–22.
Lei, Z., Wright, B.D., 2017. Why weak patents? Testing the examiner ignorance hypothesis. J. Public Econ. 148, 43–56.
Lou, B., Wu, L., 2021. AI on drugs: can artificial intelligence accelerate drug development? Evidence from a large-scale examination of bio-pharma firms. MIS Q. 45 (3), 1451–1482.
Malandraki-Miller, S., Riley, P.R., 2021. Use of artificial intelligence to enhance phenotypic drug discovery. Drug Discov. Today 26 (4), 887–901.
Manyika, S.J.P.B.J., 2019. What do we do about the biases in AI. Harv. Bus. Rev. 25.
Mikalef, P., Gupta, M., 2021. Artificial intelligence capability: conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. Inf. Manag. 58 (3), 103434.
Momeni, A., Rost, K., 2016. Identification and monitoring of possible disruptive technologies by patent-development paths and topic modeling. Technol. Forecast. Soc. Chang. 104, 16–29.
Muguerza, P., 2020. AI's infiltration of pharma: how COVID-19 accelerated change. PharmaPhorum. https://pharmaphorum.com/views-analysis-digital/ai-infiltration-ofpharma-how-covid-19-accelerated-change.
Nolan, A., 2020. Artificial intelligence, digital technology and advanced production. The Digitalisation of Science, Technology and Innovation: Key Developments.
Palangkaraya, A., Webster, E., Jensen, P.H., 2011. Misclassification between patent offices: evidence from a matched sample of patent applications. Rev. Econ. Stat. 93 (3), 1063–1075.
Pammolli, F., Magazzini, L., Riccaboni, M., 2011. The productivity crisis in pharmaceutical R&D. Nat. Rev. Drug Discov. 10 (6), 428–438.
Porter, M.E., Stern, S., 2001. Innovation: location matters. MIT Sloan Manag. Rev. 42 (4), 28.
Priem, R.L., Butler, J.E., 2001. Is the resource-based "view" a useful perspective for strategic management research? Acad. Manage. Rev. 26 (1), 22–40.
Rammer, C., Fernandez, G.P., Czarnitzki, D., 2022. Artificial intelligence and industrial′ innovation: evidence from German firm-level data. Res. Policy 51 (7), 104555.
Ransbotham, S., Gerbert, P., Reeves, M., Kiron, D., Spira, M., 2018. Artificial intelligence in business gets real. MIT Sloan Manag. Rev. 60280.
Rathi, S., Chakrabarti, A.S., Chatterjee, C., Hegde, A., 2022. Pandemics and technology engagement: new evidence from m-health intervention during COVID-19 in India. Rev. Dev. Econ. 26 (4), 2184–2217.
Rotolo, D., Hicks, D., Martin, B.R., 2015. What is an emerging technology? Res. Policy 44 (10), 1827–1843.
Samantha, M., 2020. Challenges of artificial intelligence adoption in healthcare. In: HIT Infrastructure.
Schuhmacher, A., Wilisch, L., Kuss, M., Kandelbauer, A., Hinder, M., Gassmann, O., 2021. R&D efficiency of leading pharmaceutical companies– a 20-year analysis. Drug Discov. Today 26 (8), 1784–1789.
Schumpeter, J., 1942. Capitalism, Socialism and Democracy. Harper and Row.
Secundo, G., Del Vecchio, P., Simeone, L., Schiuma, G., 2020. Creativity and stakeholders' engagement in open innovation: design for knowledge translation in technologyintensive enterprises. J. Bus. Res. 119, 272–282.
Shanbhogue, M.H., Thirumaleshwar, S., Tegginamath, P.K., Somareddy, H.K., 2021. Artificial intelligence in pharmaceutical field-a critical review. Curr. Drug Deliv. 18 (10), 1456–1466.
Singh, A., Triulzi, G., Magee, C.L., 2021. Technological improvement rate predictions for all technologies: use of patent data and an extended domain description. Res. Policy 50 (9), 104294.
Sipior, J.C., 2020. Considerations for development and use of AI in response to COVID-19. Int. J. Inf. Manag. 55, 102170.
Smalley, E., 2017. AI-powered drug discovery captures pharma interest. Nat. Biotechnol. 35 (7), 604–606.
Stempniak, M., 2022. AI adoption is lagging in radiology and the rest of healthcare. In: Radiology Business.
Tabakovic, H., Wollmann, T.G., 2018. From revolving doors to regulatory capture? Evidence from patent examiners. In: National Bureau of Economic Research.
Tietze, F., Vimalnath, P., Aristodemou, L., Molloy, J., 2020. Crisis-critical intellectual property: findings from the COVID-19 pandemic. IEEE Trans. Eng. Manag. 69 (5), 2039–2056.
Trajtenberg, M., 2018. Artificial intelligence as the next GPT: A political-economy perspective. In: The Economics of Artificial Intelligence: An Agenda. University of Chicago Press, pp. 175–186.
Wagner, S., Wakeman, S., 2016. What do patent-based measures tell us about product commercialization? Evidence from the pharmaceutical industry. Res. Policy 45 (5), 1091–1102.
Wang, Q.R., Zheng, Y., 2022. Nest without birds: inventor mobility and the left-behind patents. Res. Policy 51 (4), 104485.
Wu, L., Hitt, L., Lou, B., 2020. Data analytics, innovation, and firm productivity. Manag. Sci. 66 (5), 2017–2039.
Yamashita, I., Murakami, A., Cairns, S., Galindo-Rueda, F., 2021. Measuring the AI Content of Government-funded R&D Projects: A Proof of Concept for the OECD Fundstat Initiative. OECD.
Yang, C.-H., 2022. How artificial intelligence technology affects productivity and employment: firm-level evidence from Taiwan. Res. Policy 51 (6), 104536.
Zhou, Y., Wang, F., Tang, J., Nussinov, R., Cheng, F., 2020. Artificial intelligence in COVID-19 drug repurposing. Lancet Digit. Health 2 (12), e667–e676.
Zielinski, A., 2021. AI and the future of pharmaceutical research. In: arXiv preprint arXiv: 2107.03896.