Contents lists available at ScienceDirect

International Journal of Pharmaceutics

The role of artificial intelligence in generating original scientific research

Moe Elbadawi [*], Hanxiang Li, Abdul W. Basit, Simon Gaisford [*]

UCL School of Pharmacy, University College London, 29-39 Brunswick Square, London WC1N 1AX, UK

ABSTRACT

Artificial intelligence (AI) is a revolutionary technology that is finding wide application across numerous sectors. Large language models (LLMs) are an emerging subset technology of AI and have been developed to communicate using human languages. At their core, LLMs are trained with vast amounts of information extracted from the internet, including text and images. Their ability to create human-like, expert text in almost any subject means they are increasingly being used as an aid to presentation, particularly in scientific writing. However, we wondered whether LLMs could go further, generating original scientific research and preparing the results for publication. We tasked GPT-4, an LLM, to write an original pharmaceutics manuscript, on a topic that is itself novel. It was able to conceive a research hypothesis, define an experimental protocol, produce photo-realistic images of 3D printed tablets, generate believable analytical data from a range of instruments and write a convincing publication-ready manuscript with evidence of critical interpretation. The model achieved all this is less than 1 h. Moreover, the generated data were multi-modal in nature, including thermal analyses, vibrational spectroscopy and dissolution testing, demonstrating multi-disciplinary expertise in the LLM. One area in which the model failed, however, was in referencing to the literature. Since the generated experimental results appeared believable though, we suggest that LLMs could certainly play a role in scientific research but with human input, interpretation and data validation. We discuss the potential benefits and current bottlenecks for realising this ambition here.

1. Introduction

Artificial intelligence (AI) is a ground-breaking technology that is driving advancements in both technology and society in many fields (Briganti and Le Moine, 2020; Palagi and Fischer, 2018; Wang et al., 2022b; Wang et al., 2023b). Its primary goal is to mimic human intelligence and, as a result, to carry out human tasks (Xu et al., 2021), but at a much faster pace than humans can achieve. This capability can solve challenges like workforce shortages and eliminates the need to expose humans to hazardous situations (Gao et al., 2021). In the drug discovery process, AI provides virtual simulations, which can significantly reduce the time needed for introducing new molecules to market (Chen et al., 2018; Das et al., 2021; Popova et al., 2018). This is invaluable given the escalating cost of developing products to commercial launch. Consequently, the pharmaceutical industry has begun to explore the applications of AI to product development (Elbadawi et al., 2021).

Machine learning (ML), a branch of AI, is instrumental in increasing the efficiency of complex processes, such as forecasting three dimensional (3D) printing capabilities (Elbadawi et al., 2020,2024), predicting

drug-food interactions (Gavins et al., 2022; Kim et al., 2022), and modelling long-acting injectables (Bannigan et al., 2023). Another AI subset, machine vision (MV), is being used for tasks such as real-time monitoring of the disintegration of oral films, and is a key element in the application of process analytical technology (PAT) to tablet coating (Ficzere et al., 2022; Galata et al., 2021; O'Reilly et al., 2021; Rodrigues et al., 2021). Additionally, AI is helping the development of robotics by mimicking human movements effectively (Langer et al., 2019; von Erlach et al., 2020).

A less commonly used subset of AI, in pharmaceutics at least, is natural language processing (NLP), which aims to replicate human conversation, enhancing machine-human communication (Holler and Levinson, 2019; Trenfield et al., 2022). This allows enhanced access to machines and digital content, making the technology more accessible. Historically, interacting with machines primarily required coding, a skill not widely held. This barrier hindered researchers eager to harness the power of AI for solving pharmaceutical challenges. However, after years in development, a breakthrough in NLP was made by the development of large language models (LLMs), which has made NLP available to the

masses. These models, with access to a vast number of data, deliver on-demand intelligent responses to questions posed by human users (Agathokleous et al., 2023; De Angelis et al., 2023). This is in stark contrast to the time needed for manual retrieval of information by sifting through published work. With the overwhelming surge in scientific publications, manually locating specific information has become an arduous task. For instance, answering a seemingly simple question like "how many types of 3D printing technologies exist?" can be challenging given the expanse and breadth of the pharmaceutical literature. This has left an unmet need in the 21st century for a more efficient means of extracting relevant information (Trewartha et al., 2022). Faster information retrieval in theory should result in faster discoveries and developments.

At their core, LLMs utilise neural networks trained on billions of words and images sourced from the internet, aiming to identify connections between them (Thirunavukarasu et al., 2023). During this process, the model learns patterns, facts, grammar and even how words and ideas relate to each other. Once trained, the model can generate its own text, answer questions, or help with tasks by drawing on a large pool of learned information, almost like recalling knowledge from a gigantic digital brain. In 2020, LLMs, like the generated pre-trained transformer (GPT), could execute tasks using minimal instructions (Floridi and Chiriatti, 2020). Since then, LLM technology has advanced such that they can understand multi-modal data, like sound and visual information and its uses now encompasses generating on-demand content such as text and images. Therefore, LLMs are categorised as generative AI models, distinguishing them from earlier studies that used ML primarily for predicting outcomes.

LLMs have shown promise in generating new content, especially in the medical field, by aiding in automating written tasks. For example, Kung et al. (2023) showcased how Chat-GPT can aid in clinical decision-making. In academia, some publications have credited LLMs as co-authors, highlighting their contribution to scientific literature, and most journals now require authors to declare any use of AI. LLMs have contributed to writing review articles and even crafting experimental procedures (Frye, 2022; Marquez et al., 2023; Norris, 2023; Rahimi et al., 2023). In plant science, they have been employed to pose 'key questions in plant science' (Agathokleous et al., 2023). However, to the best of our knowledge, LLMs have not yet written a data-driven, original research article from inception to publication. In fields like pharmaceutics, creating an original, hypothesis-driven research article with accompanying data and critical interpretation is a resource-intensive endeavour requiring expertise, skill, equipment, instrumentation and materials. If LLMs can handle such a task, they could revolutionise the research landscape. This would not only illustrate their capacity for information retrieval but also their potential to produce original content, surpassing tasks like literature review writing.

To that end, we tested an LLM, GPT-4, setting it the task of writing an original data-driven pharmaceutical research paper. We asked it to create a research hypothesis, generate the accompanying data to be discussed and write a submission-ready manuscript in the authors' field of expertise; 3D printing of medicines. 3D printing is an emerging technology in manufacturing medicines and has shown great potential for addressing the lack of personalised and precise medicines (Dedeloudi et al., 2023; Elbadawi et al., 2023; Englezos et al., 2023). The technology remains in its nascent phase, and thus there are relatively few data publicly available for training LLMs. We tasked the model with imagining how a tablet comprising paracetamol dispersed in PLGA with candurin would be fabricated with selective laser sintering (SLS) printing. PLGA was selected as the main excipient because its use in pharmaceutical SLA 3D printing has not previously been evaluated (and so no literature data were available to the LLM) and it is expensive (so evaluating its use with AI potentially saves a lot of research cost). The AI-generated manuscript is appended at the end of the manuscript and we discuss our experience with GPT-4 in this report.

2. Experimental procedure

Prompts (full details of which are appended in the supplementary information) were submitted to ChatGPT PLUS using the GPT-4 model (Chat-GPT July 20 Version). Text responses from prompts were copied and pasted into Word. To generate data, the model was prompted to generate representative code for python. The code was then ported to python (v3.11.1) and the plots generated were transferred to the manuscript. For image production, the application programming interface (API) supplied by OpenAI for python was used (openai v.0.27.8). Image generation was performed with the DALL-E model, which is a generative NLP model. Similar to GPT-4, the model was prompted with text and the generated images were copied and pasted into the manuscript.

3. Results and discussion

GPT-4′s ability to write an original manuscript, including within the text data, plots and images, was nothing short of remarkable. In less than one hour, the AI platform was able to generate the results of a study and prepare a manuscript for publication. In other words, the entire contents of the manuscript presented below were generated *de novo* by AI. While it may not be surprising that GPT-4 could generate the words to describe the data, the fact that it could generate data to support a research hypothesis is a key, and somewhat unexpected, finding.

Typically, users interact with GPT-4 through text, but our study has shown that LLMs can go further, and convert text into data, figures and images, all of which are common means of data representation in manuscripts. We chose to prompt the LLM with a drug and polymer combination (paracetamol with PLGA and a colourant dye, candurin) which had not been previously reported in the literature; hence, it was not possible for the LLM to simply retrieve data from the internet. It had to create data and images *de novo*.

Different types of data, including spectroscopic, optical and x-ray micro-computed tomography (XRMCT), were created and they looked compelling. Additionally, the model provided a critical commentary of the data. For example, it produced sensible glass transition and melting temperatures for PLGA and a melting temperature for paracetamol, and knew how these would manifest in a differential scanning calorimetry (DSC) thermogram (Table 1) (Lanao et al., 2013). Similarly, it produced prototypical degradation plots for these components and was able to simulate their thermal gravimetric analysis (TGA) curves (Awad et al., 2019; Giri and Maniruzzaman, 2022; Shi et al., 2018; Zhang et al., 2023). For human researchers, this level of knowledge retrieval would

Table 1

Features of the simulated TGA and DSC compared to real-world examples.

Characterisation technique	Material	Feature	Reference
TGA	PLGA	Degradation onset ~300 °C, and complete degradation by 350 °C	(Jose et al., 2009)
	Paracetamol	Degradation onset ~250 °C and complete degradation by 300 °C	(de Oliveira et al., 2017; Goyanes et al., 2015)
	Candurin	Thermally stable until 500 °C, with minor weight loss	(Zhang et al., 2023)
DSC	PLGA	T_g at ~60 °C and T_m ~ 140 °C	(Walejewska et al., 2020)
	Paracetamol	T_m ~ 170 °C	(Khaled et al., 2018b)
	Candurin	No thermal events between 25 and 250 °C.	(Madžarević et al., 2021)

require an exhaustive period of literature surveying, for each component, and matching that with experimental data.

The simulated Fourier-transformed infrared (FTIR) spectra were particularly good. FTIR data typically require multi-variate analysis to interpret and is not a trivial task. Here, AI was able to work through it logically. First, it recalled the chemical structure of, for example, paracetamol. Thereafter, it postulated potential vibration bands based on the chemical structure of the material and how they would manifest themselves in an FTIR plot. As a result, the simulated FTIR plots were indistinguishable from real FTIR plots (Table 2). For XRD, GPT-4 was able to classify the materials as either crystalline, semi-crystalline or amorphous and was able to produce intensity peaks for each material (Table 2).

While the ability of the model to generate data was good, its ability to demonstrate critical thinking was even more impressive. For instance, it postulated an effect of laser scanning speed on the mechanical and dissolution properties of the printlets, even though there is no template or previous precedence for this in the literature. It also showed a relationship between laser speed and printlet porosity, and used this to explain differences in mechanical properties and dissolution profiles. Interestingly, its ability to analyse critically how paracetamol might become amorphous during SLS printing, and how this might alter DSC, FTIR and X-ray diffraction (XRD) data, demonstrate a fundamental understanding of both material and pharmaceutical sciences.

We also used AI's text-to-image feature to generate images of the printlets. While the main object (i.e. the printlet) was perfectly captured, the model can be seen to struggle with the surrounding content. In particular, the ruler hash marks can be seen to be abnormal. Minor abnormalities in photographic images like this are a key indicator that they are not real, and this may be an important tool in determining whether images are real or artificial.

The model was also tasked with generating X-ray Micro Computed Tomography (XRMCT) images. Here, the images showed printlets with a similar morphology to previous work (Fina et al., 2017), although improvements would be needed to make these images appear more realistic.

Finally, the ability of the model to write was also striking. Communicating the results of a study is a critical aspect of scientific research, and a lot of information can be embedded in text that cannot be otherwise communicated. In addition to the results section, GPT-4 was able to generate a methodology section, including within it some very detailed experimental protocols, similar to those noted by Marquez et al., (2023). The ability to create such detailed work plans may help guide researchers who are new to a particular discipline or technique, especially in today's cross-disciplinary environment. The model was also able to rationalise the need for the study and provided some background information in the Introduction.

There were two areas where the model showed any deficiencies. One was a lack of keywords, and the other was in referencing literature. Indeed, no references were cited and it is not clear why the model was unable to accomplish this. It may be that the model seeks data from the internet, assuming all information is not attributable to specific authors, and does not scan individual research papers, in the way that a human researcher would. Of course, many of the data the model generated were created de novo, and so could not be cited, but the lack of citations in the introduction section is a clear weakness, although this may be used as the basis of a method for identifying text that has been generated with AI.

Overall though, by showing an ability to write an original research article, AI has achieved a significant breakthrough in simulating human intelligence. The results of this study suggest that LLMs have the potential to transform pharmaceutical research radically, despite their infancy. The authors have been interested in using AI to automate aspects of the pharmaceutical research pipeline in the interest of accelerating discoveries and developments and doing so in an environmentally sustainable manner (Abdalla et al., 2023; McCoubrey et al., 2022; Wang et al., 2023a) and we have been successful in modelling and automating many aspects of the research pipeline. However, we have always needed the laborious steps of data collection and pre-processing of information to feed into an AI model. Here, in contrast, no data collection or data pre-processing was needed; The LLM generated everything *de novo*, which allowed completion of its task at a fast pace.

We have not yet experimentally validated the outcomes of the model, by printing and characterising PLGA/paracetamol tablets but what has been achieved within this study builds on our previous work. For example, AI's ability to simulate FTIR data from simple text prompts is unprecedented and opens up new avenues of sustainable simulations and the prospect of simulating the entire research pipeline appears feasible. Further 'stress tests' are needed to see how LLMs can cope with human inter- and intra-variability, which are known to cause variability in data and has been an attributor of data irreproducibility. Indeed, other sources of variations, such as ambient temperature and humidity variation, should be factored in by LLM when generating simulated data, and it will be interesting to see how the platform can adapt to these unpredictable scenarios, as well as being integrated into Internet of Things (IoT) framework (Olvera and Monaghan, 2021; Raijada et al., 2021).

Its ability to write a manuscript on a research topic that itself is emerging was incredible. Relative to other pharmaceutical research topics, there is limited information surrounding SLS printing of medicines (Charoo et al., 2020). It is anticipated that as the knowledge of SLS develops, so too will AI's prowess of the topic. SLS printing of PLGA was selected because it has not been published nor documented and is of personal interest to the authors. Additionally, PLGA is expensive and so conducting experimental research with it requires significant funding. This work suggests that LLMs could be used to predict the outcomes of using expensive materials in research, and the results could be used to select which materials are used for real studies. In hindsight, PLGA was an ideal polymer for this study due to the copious amount of information available on its use and because of its applications in many material and healthcare sectors (Wang et al., 2022a). This is in contrast to some pharmaceutical polymers for which there are a lack of published data because they are almost exclusively used in pharmaceutical research and their chemical structures have not been disclosed by their manufacturers. It would be interesting to see how AI would simulate data based on these materials.

Data remains the main issue in using AI in pharmaceutics. All AI systems look to published data to draw relationships between chemical structure, physicochemical properties and behaviour in formulated

Table 2

Features of the simulated FTIR and XRD compared to real-world examples.

Characterisation technique	Material	Feature	Reference
FTIR	PLGA	Characteristic single peak at 1750 cm^{-1} and multiple peaks between 1550 and 850 cm^{-1}	(Dou et al., 2021; Wei et al., 2022)
	Paracetamol	Characteristic band ~3200 cm^{-1} and multiple peaks between 1600 and 500 cm^{-1}	(Khaled et al., 2018a)
	Candurin	Characteristic peak ~1000 cm^{-1}	(Zhang et al., 2023)
XRD	PLGA	Semi-crystalline; few peaks	(Jeong et al., 2023)
	Paracetamol	Crystalline; multiple peaks	(Khaled et al., 2018a; Prasad et al., 2019)
	Candurin	Crystalline; few peaks	(Davis et al., 2020)

medicines. Without open source data an AI system cannot develop relationships which it can use to predict outcomes. On the other hand, because the use of AI to generate scientific data is a new paradigm, most of the data in the literature have been generated by experimental research and any relationships between materials is real. If AI-generated data begin to populate the internet, then there will be an increasing proportion of data that are not real, and there is a risk that AI models start to predict non-sensical outcomes (this is already an issue being seen in the field of art, for instance). It may be the case that technologies such as blockchain can circumvent this issue (Trenfield et al., 2022), but the authors strongly suggest that all published data generated with AI are marked as such, so that they are not incorporated into future predictions by AI models. We also note that regulations will indeed be needed. For one, models should be closely monitored to ensure that they are trained with high quality, unbiased data, and they should be robust to adversarial attacks (Chen et al., 2023; Kaviani et al., 2022). There has been concern regarding LLMs 'hallucinating' responses, whereby they generate fictitious information (Brodnik et al., 2023). However, this is being actively addressed and once achieved, it is anticipated that it will result in more accurate experimental simulations.

The ability of LLMs to generate different data of multiple types clearly demonstrates multi-disciplinary expertise beyond the pharmaceutical sciences. Future work will seek to stretch its use to new data modalities and to evaluate the extent of its multi-disciplinary expertise. In addition, while communicating with AI via human languages makes it more widely accessible than communicating with it via coding, it will be interesting to see if it can be made even more accessible, for example by ensuring any AI platform will be economically viable and not hidden behind a paywall (Liebrenz et al., 2023).

4. Conclusion

We have demonstrated how GPT-4, an LLM, can simulate completion of a research project on a topic that is itself novel. It was able to conceive a research hypothesis, define an experimental protocol, produce photo-realistic images of the printlets, generate believable analytical data from a range of instruments and write a convincing publication-ready manuscript with evidence of critical interpretation. The model achieved all this is less than 1 h. While caution must be exercised in the value placed on the research outcomes, we have nonetheless shown the potential power of AI in accelerating research. If the data generated this way are representative of reality, then AI could be used to save time and cost as well as limit the environmental impact of research.

CRediT authorship contribution statement

Moe Elbadawi: Conceptualization, Data curation, Formal analysis, Writing – original draft, Writing – review & editing. **Hanxiang Li:** . **Abdul W. Basit:** . **Simon Gaisford:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijpharm.2023.123741.

References

Abdalla, Y., Elbadawi, M., Ji, M., Alkahtani, M., Awad, A., Orlu, M., Gaisford, S., Basit, A. W., 2023. Machine learning using multi-modal data predicts the production of selective laser sintered 3D printed drug products. Int. J. Pharm. 633, 122628.

Agathokleous, E., Rillig, M.C., Peñuelas, J., Yu, Z., 2023. One hundred important questions facing plant science derived using a large language model. Trends Plant Sci.

Awad, A., Fina, F., Trenfield, S.J., Patel, P., Goyanes, A., Gaisford, S., Basit, A.W., 2019. 3D printed pellets (Miniprintlets): a novel, multi-drug, controlled release platform technology. Pharmaceutics 11, 148.

Bannigan, P., Bao, Z., Hickman, R.J., Aldeghi, M., Häse, F., Aspuru-Guzik, A., Allen, C., 2023. Machine learning models to accelerate the design of polymeric long-acting injectables. Nat. Commun. 14, 35.

Briganti, G., Le Moine, O., 2020. Artificial intelligence in medicine: today and tomorrow. Front. Med. 7.

Brodnik, N.R., Carton, S., Muir, C., Ghosh, S., Downey, D., Echlin, M.P., Pollock, T.M., Daly, S., 2023. Perspective: large language models in applied mechanics. J. Appl. Mech. 90.

Charoo, N.A., Barakh Ali, S.F., Mohamed, E.M., Kuttolamadom, M.A., Ozkan, T., Khan, M.A., Rahman, Z., 2020. Selective laser sintering 3D printing – an overview of the technology and pharmaceutical applications. Drug Dev. Ind. Pharm. 46, 869–877.

Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., Blaschke, T., 2018. The rise of deep learning in drug discovery. Drug Discov. Today 23, 1241–1250.

Chen, R.J., Wang, J.J., Williamson, D.F.K., Chen, T.Y., Lipkova, J., Lu, M.Y., Sahai, S., Mahmood, F., 2023. Algorithmic fairness in artificial intelligence for medicine and healthcare. Nat. Biomed. Eng. 7, 719–742.

Das, P., Sercu, T., Wadhawan, K., Padhi, I., Gehrmann, S., Cipcigan, F., Chenthamarakshan, V., Strobelt, H., dos Santos, C., Chen, P.-Y., Yang, Y.Y., Tan, J.P. K., Hedrick, J., Crain, J., Mojsilovic, A., 2021. Accelerated antimicrobial discovery via deep generative models and molecular dynamics simulations. Nat. Biomed. Eng. 5, 613–623.

Davis, D.A., Miller, D.A., Su, Y., Williams, R.O., 2020. Thermally conductive excipient expands kinetisol® processing capabilities. AAPS PharmSciTech 21, 319.

De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G.P., Ferragina, P., Tozzi, A.E., Rizzo, C., 2023. ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. Front. Public Health 11.

de Oliveira, G.G.G., Feitosa, A., Loureiro, K., Fernandes, A.R., Souto, E.B., Severino, P., 2017. Compatibility study of paracetamol, chlorpheniramine maleate and phenylephrine hydrochloride in physical mixtures. Saudi Pharmaceut. J. 25, 99–103.

Dedeloudi, A., Weaver, E., Lamprou, D.A., 2023. Machine learning in additive manufacturing & microfluidics for smarter and safer drug delivery systems. Int. J. Pharm. 636, 122818.

Dou, Y., Huang, J., Xia, X., Wei, J., Zou, Q., Zuo, Y., Li, J., Li, Y., 2021. A hierarchical scaffold with a highly pore-interconnective 3D printed PLGA/n-HA framework and an extracellular matrix like gelatin network filler for bone regeneration. J. Mater. Chem. B 9, 4488–4501.

Elbadawi, M., Muñiz Castro, B., Gavins, F.K.H., Ong, J.J., Gaisford, S., Pérez, G., Basit, A. W., Cabalar, P., Goyanes, A., 2020. M3DISEEN: a novel machine learning approach for predicting the 3D printability of medicines. Int. J. Pharm. 590, 119837.

Elbadawi, M., McCoubrey, L.E., Gavins, F.K.H., Ong, J.J., Goyanes, A., Gaisford, S., Basit, A.W., 2021. Disrupting 3D printing of medicines with machine learning. Trends Pharmacol. Sci. 42, 745–757.

Elbadawi, M., Basit, A.W., Gaisford, S., 2023. Energy consumption and carbon footprint of 3D printing in pharmaceutical manufacture. Int. J. Pharm. 639, 122926.

Elbadawi, M., Li, H., Basit, A., Gaisford, S., 2024. Artificial Intelligence generates novel 3D printing formulations. Appl. Mater. Today. https://doi.org/10.1016/j.apmt.2024.102061.

Englezos, K., Wang, L., Tan, E.C.K., Kang, L., 2023. 3D printing for personalised medicines: implications for policy and practice. Int. J. Pharm. 635, 122785.

Ficzere, M., Mészáros, L.A., Kállai-Szabó, N., Kovács, A., Antal, I., Nagy, Z.K., Galata, D. L., 2022. Real-time coating thickness measurement and defect recognition of film coated tablets with machine vision and deep learning. Int. J. Pharm. 623, 121957.

Fina, F., Goyanes, A., Gaisford, S., Basit, A.W., 2017. Selective laser sintering (SLS) 3D printing of medicines. Int. J. Pharm. 529, 285–293.

Floridi, L., Chiriatti, M., 2020. GPT-3: its nature, scope, limits, and consequences. Mind. Mach. 30, 681–694.

Frye, B.L., 2022. Should using an AI text generator to produce academic writing be plagiarism? Fordham Intellectual Property, Media Entertain. Law J., Forthcoming.

Galata, D.L., Mészáros, L.A., Kállai-Szabó, N., Szabó, E., Pataki, H., Marosi, G., Nagy, Z. K., 2021. Applications of machine vision in pharmaceutical technology: a review. Eur. J. Pharm. Sci. 159, 105717.

Gao, A., Murphy, R.R., Chen, W., Dagnino, G., Fischer, P., Gutierrez, M.G., Kundrat, D., Nelson, B.J., Shamsudhin, N., Su, H., Xia, J., Zemmar, A., Zhang, D., Wang, C., Yang, G.-Z., 2021. Progress in robotics for combating infectious diseases. Sci. Robot. 6, eabf1462.

Gavins, F.K.H., Fu, Z., Elbadawi, M., Basit, A.W., Rodrigues, M.R.D., Orlu, M., 2022. Machine learning predicts the effect of food on orally administered medicines. Int. J. Pharm. 611, 121329.

Giri, B.R., Maniruzzaman, M., 2022. Fabrication of sustained-release dosages using powder-based three-dimensional (3D) printing technology. AAPS PharmSciTech 24, 4.

Goyanes, A., Robles Martinez, P., Buanz, A., Basit, A.W., Gaisford, S., 2015. Effect of geometry on drug release from 3D printed tablets. Int. J. Pharm. 494, 657–663.

Holler, J., Levinson, S.C., 2019. Multimodal language processing in human communication. Trends Cogn. Sci. 23, 639–652.

Jeong, J., Yoon, S., Yang, X., Kim, Y.J., 2023. Super-tough and biodegradable poly (lactide-co-glycolide) (PLGA) transparent thin films toughened by star-shaped PCL-b-PDLA plasticizers. Polymers 15, 2617.

Jose, M.V., Thomas, V., Dean, D.R., Nyairo, E., 2009. Fabrication and characterization of aligned nanofibrous PLGA/Collagen blends as bone tissue scaffolds. Polymer 50, 3778–3785.

Kaviani, S., Han, K.J., Sohn, I., 2022. Adversarial attacks and defenses on AI in medical imaging informatics: a survey. Exp. Syst. Appl. 198, 116815.

Khaled, S.A., Alexander, M.R., Irvine, D.J., Wildman, R.D., Wallace, M.J., Sharpe, S., Yoo, J., Roberts, C.J., 2018a. Extrusion 3D printing of paracetamol tablets from a single formulation with tunable release profiles through control of tablet geometry. AAPS PharmSciTech 19, 3403–3413.

Khaled, S.A., Alexander, M.R., Wildman, R.D., Wallace, M.J., Sharpe, S., Yoo, J., Roberts, C.J., 2018b. 3D extrusion printing of high drug loading immediate release paracetamol tablets. Int. J. Pharm. 538, 223–230.

Kim, S., Choi, Y., Won, J.-H., Mi Oh, J., Lee, H., 2022. An annotated corpus from biomedical articles to construct a drug-food interaction database. J. Biomed. Inform. 126, 103985.

Kung, T.H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., Tseng, V., 2023. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. PLOS Digital Health 2, e0000198.

Lanao, R.P.F., Jonker, A.M., Wolke, J.G., Jansen, J.A., van Hest, J.C., Leeuwenburgh, S.C., 2013. Physicochemical properties and applications of poly (lactic-co-glycolic acid) for use in bone regeneration. Tissue Eng. B Rev. 19, 380–390.

Liebrenz, M., Schleifer, R., Buadze, A., Bhugra, D., Smith, A., 2023. Generating scholarly content with ChatGPT: ethical challenges for medical publishing. Lancet Dig. Health 5, e105–e106.

Madžarević, M., Medarević, Đ., Pavlović, S., Ivković, B., Đuriš, J., Ibrić, S., 2021. Understanding the effect of energy density and formulation factors on the printability and characteristics of SLS irbesartan tablets—application of the decision tree model. Pharmaceutics 13, 1969.

Marquez, R., Barrios, N., Vera, R.E., Mendez, M.E., Tolosa, L., Zambrano, F., Li, Y., 2023. A perspective on the synergistic potential of artificial intelligence and product-based learning strategies in biobased materials education. Educ. Chem. Eng. 44, 164–180.

McCoubrey, L.E., Seegobin, N., Elbadawi, M., Hu, Y., Orlu, M., Gaisford, S., Basit, A.W., 2022. Active machine learning for formulation of precision probiotics. Int. J. Pharm. 616, 121568.

Norris, C., 2023. Large language models like ChatGPT in ABME: author guidelines. Ann. Biomed. Eng. 51, 1121–1122.

O'Reilly, C.S., Elbadawi, M., Desai, N., Gaisford, S., Basit, A.W., Orlu, M., 2021. Machine learning and machine vision accelerate 3D printed orodispersible film development. Pharmaceutics 13, 2187.

Olvera, D., Monaghan, M.G., 2021. Electroactive material-based biosensors for detection and drug delivery. Adv. Drug Deliv. Rev. 170, 396–424.

Palagi, S., Fischer, P., 2018. Bioinspired microrobots. Nat. Rev. Mater. 3, 113–124.

Popova, M., Isayev, O., Tropsha, A., 2018. Deep reinforcement learning for de novo drug design. Sci. Adv. 4, eaap7885.

Prasad, E., Islam, M.T., Goodwin, D.J., Megarry, A.J., Halbert, G.W., Florence, A.J., Robertson, J., 2019. Development of a hot-melt extrusion (HME) process to produce drug loaded Affinisol™ 15LV filaments for fused filament fabrication (FFF) 3D printing. Addit. Manuf. 29, 100776.

Rahimi, F., Talebi Bezmin Abadi, A., 2023. ChatGPT and publication ethics. Arch. Med. Res. 54, 272–274.

Raijada, D., Wac, K., Greisen, E., Rantanen, J., Genina, N., 2021. Integration of personalized drug delivery systems into digital health. Adv. Drug Deliv. Rev. 113857.

Rodrigues, C.P., Duchesne, C., Poulin, É., Lapointe-Garant, P.-P., 2021. In-line cosmetic end-point detection of batch coating processes for colored tablets using multivariate image analysis. Int. J. Pharm. 606, 120953.

Shi, Y., Xue, J., Jia, L., Du, Q., Niu, J., Zhang, D., 2018. Surface-modified PLGA nanoparticles with chitosan for oral delivery of tolbutamide. Colloids Surf. B Biointerf. 161, 67–72.

Thirunavukarasu, A.J., Ting, D.S.J., Elangovan, K., Gutierrez, L., Tan, T.F., Ting, D.S.W., 2023. Large language models in medicine. Nat. Med.

Trenfield, S.J., Awad, A., McCoubrey, L.E., Elbadawi, M., Goyanes, A., Gaisford, S., Basit, A.W., 2022. Advancing pharmacy and healthcare with virtual digital technologies. Adv. Drug Deliv. Rev. 182, 114098.

Trewartha, A., Walker, N., Huo, H., Lee, S., Cruse, K., Dagdelen, J., Dunn, A., Persson, K.A., Ceder, G., Jain, A., 2022. Quantifying the advantage of domain-specific pre-training on named entity recognition tasks in materials science. Patterns 3, 100488.

von Erlach, T., Saxton, S., Shi, Y., Minahan, D., Reker, D., Javid, F., Lee, Y.-A.-L., Schoellhammer, C., Esfandiary, T., Cleveland, C., Booth, L., Lin, J., Levy, H., Blackburn, S., Hayward, A., Langer, R., Traverso, G., 2020. Robotically handled whole-tissue culture system for the screening of oral drug formulations. Nat. Biomed. Eng. 4, 544–559.

Walejewska, E., Idaszek, J., Heljak, M., Chlanda, A., Choinska, E., Hasirci, V., Swieszkowski, W., 2020. The effect of introduction of filament shift on degradation behaviour of PLGA- and PLCL-based scaffolds fabricated via additive manufacturing. Polym. Degrad. Stab. 171, 109030.

Wang, F., Elbadawi, M., Tsilova, S.L., Gaisford, S., Basit, A.W., Parhizkar, M., 2022a. Machine learning predicts electrospray particle size. Mater. Des. 219, 110735.

Wang, F., Elbadawi, M., Tsilova, S.L., Gaisford, S., Basit, A.W., Parhizkar, M., 2022b. Machine learning to empower electrohydrodynamic processing. Mater. Sci. Eng. C 132, 112553.

Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., Chandak, P., Liu, S., Van Katwyk, P., Deac, A., Anandkumar, A., Bergen, K., Gomes, C.P., Ho, S., Kohli, P., Lasenby, J., Leskovec, J., Liu, T.-Y., Manrai, A., Marks, D., Ramsundar, B., Song, L., Sun, J., Tang, J., Veličković, P., Welling, M., Zhang, L., Coley, C.W., Bengio, Y., Zitnik, M., 2023b. Scientific discovery in the age of artificial intelligence. Nature 620, 47–60.

Wang, F., Sangfuang, N., McCoubrey, L.E., Yadav, V., Elbadawi, M., Orlu, M., Gaisford, S., Basit, A.W., 2023a. Advancing oral delivery of biologics: machine learning predicts peptide stability in the gastrointestinal tract. Int. J. Pharm. 634, 122643.

Wei, J., Yan, Y., Gao, J., Li, Y., Wang, R., Wang, J., Zou, Q., Zuo, Y., Zhu, M., Li, J., 2022. 3D-printed hydroxyapatite microspheres reinforced PLGA scaffolds for bone regeneration. Biomater. Adv. 133, 112618.

Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., Qiu, C.-W., Qiu, J., Hua, K., Su, W., Wu, J., Xu, H., Han, Y., Fu, C., Yin, Z., Liu, M., Roepman, R., Dietmann, S., Virta, M., Kengara, F., Zhang, Z., Zhang, L., Zhao, T., Dai, J., Yang, J., Lan, L., Luo, M., Liu, Z., An, T., Zhang, B., He, X., Cong, S., Liu, X., Zhang, W., Lewis, J.P., Tiedje, J.M., Wang, Q., An, Z., Wang, F., Zhang, L., Huang, T., Lu, C., Cai, Z., Wang, F., Zhang, J., 2021. Artificial intelligence: a powerful paradigm for scientific research. The Innovation 2, 100179.

Zhang, Y., Thakkar, R., Zhang, J., Lu, A., Duggal, I., Pillai, A., Wang, J., Aghda, N.H., Maniruzzaman, M., 2023. Investigating the use of magnetic nanoparticles as alternative sintering agents in selective laser sintering (SLS) 3D printing of oral tablets. ACS Biomater Sci. Eng. 9, 2924–2936.