

The perils and promises of generative artificial intelligence in neurointerventional surgery

Tyler R Ray,¹ Ryan T Kellogg ,² Kyle M Fargen ,³ Ferdinand Hui ,⁴ Jan Vargas ⁵

SUMMARY

Generative artificial intelligence (AI) holds great promise in neurointerventional surgery by providing clinicians with powerful tools for improving surgical precision, accuracy of diagnoses, and treatment planning. However, potential perils include biases or inaccuracies in the data used to train the algorithms, over-reliance on generative AI without human oversight, patient privacy concerns, and ethical implications of using AI in medical decision-making. Careful regulation and oversight are needed to ensure that the promises of generative AI in neurointerventional surgery are realized while minimizing its potential perils.

[ChatGPT authored summary using the prompt "In one paragraph summarize the promises and perils of generative AI in neurointerventional surgery".]

INTRODUCTION

Relatively few technological breakthroughs have arisen with the capacity to alter society as swiftly and profoundly as the printing press, penicillin, and the atomic bomb. Originating from disparate fields of inquiry, innovations such as these have left an indelible mark on our collective experience, reshaping the fabric of human history and sparking both a sense of wonder and fear on a societal level. The recent emergence of generative artificial intelligence (AI) marks another such inflection point in the course of human progress. With the potential to rapidly reshape society, this technology presents

both the prospect of transformational benefits and the specter of profound risks.

In late November 2022, the Open AI Research Laboratory quietly introduced an enhanced AI chatbot named ChatGPT to the general public.¹ In a remarkable testament to the technological sophistication and societal interest, over 100 million users subscribed to ChatGPT within the first 2 months of its release, setting a new record for any consumer internet service.² The extraordinary speed of subscriber growth for ChatGPT caught many tech industry leaders by surprise, prompting many to develop their own chatbot offerings, or to announce plans to do so in the near future.

The rapid proliferation of these services has forced society to confront significant ethical and societal questions related to generative AI, particularly its ability to emulate human thought processes and creativity. As Alan Turing eloquently captured in 1950, the question of "Can machines think?" has emerged as a pivotal philosophical query in the context of this technology.³ By definition, the development of machines that can generate creative and novel output pushes the boundaries of what we consider to be 'thinking' and 'intelligence'. The implications of such technology, particularly in the context of medicine and neurointerventional surgery, call for thoughtful reflection, dialogue, and collaboration across diverse fields of expertise to chart a course that balances the potential benefits and risks. To do so will provide a forum for addressing concerns around the accuracy of responses, the potential for generative AI to spread misinformation or manipulate human behavior, and the impact on human employment and social interaction.

GENERATIVE AI: A BRIEF PRIMER

As the field of machine learning continues to advance at a rapid pace, generative AI has emerged as a particularly dynamic subfield which focuses on the creation of new content such as images, videos, or text using algorithms, statistical models, and

extraordinarily large sets of data. One of the most powerful and popular techniques within generative AI is the use of Generative Adversarial Networks (GANs).⁴ GAN models consist of two neural networks: a generator and a discriminator. The generator produces new data samples that mimic the input data, while the discriminator distinguishes between the real and generated data. These two components compete with one another, resulting in image generation models that can create high resolution synthetic digital images (for instance, of head CT images which simply do not exist; [figure 1](#)), as with the case of StyleGAN-3 (NVIDIA, Santa Clara, California, USA).⁵

More recently, large language models such as the Generative Pre-trained Transformer (GPT) have also emerged as a significant development in the field of generative AI.⁶ These models leverage deep learning methods to generate human-like language, such as written text or speech, with remarkable accuracy and fluency, to the point that ChatGPT can almost pass the United States Medical Licensing Exam and the neurosurgical oral boards.^{7,8} GPT-3, the underlying algorithm on which ChatGPT is based, was trained on a massive corpus of text data, including a diverse range of books, articles, and websites, allowing it to learn the patterns and relationships between words, phrases, and sentences. This vast amount of data enables a model to learn a general understanding of language (eg, the patterns and relationships between words, phrases, and sentences), which can then be fine-tuned for specific tasks such as language translation or question answering by training the model on smaller, task-specific datasets. This approach can be applied to multimodal datasets (ie, text-image pairs) as demonstrated by DALL-E and DALL-E 2, which used a version of the GPT-3 algorithm to create a model that generates images from natural language text ([figure 2](#)).⁹

USE OF AI IN NEUROINTERVENTIONAL SURGERY

AI has already had a profound impact on many specialties in medicine. In neurointerventional surgery, deep learning models have been developed to detect large vessel occlusions and cerebral aneurysms on vessel imaging, leading to faster times to treatment and more streamlined workflows.^{10,11} In these cases of detection and triage, the role of the algorithm is to flag concerning studies for human review and *not* to replace the physician assessment. Volumetric measurements for several disease states can be done rapidly

¹Department of Mechanical Engineering, University of Hawaii at Mānoa College of Engineering, Honolulu, Hawaii, USA

²Department of Neurosurgery, University of Virginia, Charlottesville, Virginia, USA

³Department of Neurological Surgery and Radiology, Wake Forest University, Winston-Salem, North Carolina, USA

⁴Neurointerventional Surgery, Queen's Medical Center Neuroscience Institute, Honolulu, Hawaii, USA

⁵Division of Neurosurgery, Prisma Health Upstate, Greenville, South Carolina, USA

Correspondence to Dr Jan Vargas, Surgery, Prisma Health Upstate, Greenville, 29605, South Carolina, USA; jvargas.machaj@prismahealth.org

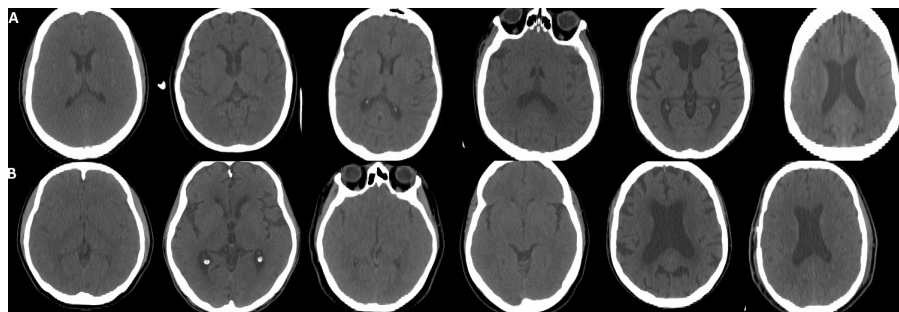


Figure 1 (A) Synthetic head CT images generated using a StyleGAN³ model trained on the personal computer of the senior author demonstrating both the relative ease and general fidelity of generated AI medical data. (B) Real patient head CT images.

and consistently, removing the burden of manual estimation from humans and paving the way for deeper understanding of cerebrovascular disorders. Volumetric endpoints for evacuation of intracranial hemorrhages, longitudinal monitoring of ventricular volumes, and measuring the volume of pathologies such as subdural collections before and after intervention have all been made possible through the use of machine learning algorithms.¹²⁻¹⁴

Automatically generated ASPECT scores have been shown to be as good or better than experienced radiologists.¹⁵ Yet even with these advances, AI in neurointerventional surgery is in its infancy, with several promising applications.



Figure 2 Panel of AI-generated images by the authors using Midjourney with the instructions "A highly detailed photo of Tyler R PhD, Ryan K MD, Kyle F MD, Ferdinand H MD, and Jan V MD standing in front of a hospital". The people pictured in the images bear no resemblance to the authors.

PROMISE OF GENERATIVE AI FOR MEDICAL APPLICATIONS

The potential to incorporate generative AI into neurointerventional surgery is vast. As observed in other fields, this technology could yield fresh and valuable insights into medical datasets through machine learning techniques.¹⁶ For instance, one area that could benefit immensely from generative models is natural language processing (NLP). These models could be used to translate clinical information into layperson-friendly formats, enhancing effective communication between medical practitioners and patients.¹⁷ As an example, consider a patient diagnosed with a new cerebral aneurysm. NLP models could rapidly provide patients with an education-level appropriate description of personal natural history risk based on best available high quality data coupled with personalized risk factors (size and location of aneurysm, smoking status, family history, medical history). Next, such a model would collate data on treatment options (clipping, coiling, stent assisted coiling, flow diversion, etc) and compile the risk versus benefits associated with each option based on automatically generated systematic review-quality data. Finally, the model would generate a list of local "preferred" providers based on research productivity, international reputation, years of independent practice, patient reviews, and institutional quality scores. Follow-up patient visits could be tailored to specific issues and questions based on previous visits and questions they had asked.

It is noteworthy that transformer models, a specific type of machine learning algorithm on which ChatGPT is based, were initially used for translating from one language to another and, today, we can envisage models that produce coherent clinical reports based on a patient's medical data. These models can operate in a mode that generates scalable explanations to patients with varying levels of education and cognitive ability or generate discharge summaries by parsing through medical records. They could also help develop more efficient billing and coding, or serve as a human-like agent that can interact with and triage patients with precision and efficiency. One area where a triaging agent could be particularly impactful would be in directing patients with convincing signs and symptoms of a large vessel occlusion to primary or comprehensive stroke centers for immediate treatment. In such cases, a prehospital assessment could be performed by a model and, if there is a high enough concern for a stroke, initiate

a telehealth visit with an available stroke neurologist while routing the patient to the appropriate center for directed treatment. In such a fast-paced situation, clear communication with the patient and his or her family is vital. Again, in such instances the role of AI would be not to replace the physician assessment, but to flag high-yield cases for immediate review.

Generative models have also been applied to medical imaging. Super resolution (SR) GANs have been used to transform lower resolution imaging data from CT and MRI scans to higher resolution.^{18–20} Such SR GANs could also be applied to angiography, allowing for upscaling and providing higher resolutions than are currently available. Other possibilities include generating synthetic datasets for medical trainees, which are no longer confined to the realm of science fiction. Trainees could use a model to produce a rendering (figure 1) of a patient's brain based on their medical history as a visual aid during their studies. In the field of device development, these models could be used to create flow models with limitless (but realistic) configurations for testing and cataloguing the behavior of new devices. Finally, recent work by Tang *et al* has focused on the use of semantic decoder networks, similar to ChatGPT, which can translate a person's thoughts into text.²¹ Such technology could be vital for stroke rehabilitation and prosthesis development, paving the way for further work with brain–computer interfaces.

By holding up the mirror to something that used to be inherently human prior to these advances, generative AI has reinforced the desire to understand the processes behind creativity, both by humans and due to generative models. For instance, further research into understanding latent space (an abstract multi-dimensional space containing features that we cannot interpret directly) from which GANs sample input can lead to more control over the final output.²² In an effort to understand how machine learning models arrive at their predictions, Shapley values have been used to investigate which inputs into a model contribute the most.²³ Although there is progress in understanding the historical black box, the complexity and novelty of these models can lead to abuse.

PERILS OF GENERATIVE AI FOR MEDICAL APPLICATIONS

The complexity of these models can also make it challenging to identify and address potential dangers that may arise from their use. Artificial neural networks are

a critical element of most deep learning models, but they have been criticised as ‘black boxes’ due to their inner workings being beyond the current understanding of researchers.²⁴ More broadly, the potential for malfeasance enabled by ChatGPT or other generative algorithms is a growing concern with the following risks already identified:

1. *Deepfake imaging*: Deepfake technology, combined with GAN, could be used to generate fake patient imaging data by blending real imaging with imaging of pathologic samples or in falsifying identities of individuals.²⁵ Mirsky *et al* have presented their work with using GANs to inject and remove lung cancer lesions from CT imaging.²⁶
2. *Falsified clinical reports*: Patients with Munchausen Syndrome or Munchausen by Proxy could use GAN and generative chatbots to create clinical reports that support a false medical narrative, which could be misused by doctors or insurance companies.²⁷
3. *Malpractice and falsified data*: The use of falsified data in malpractice lawsuits could lead to either the strengthening or defending of a case.²⁸
4. *Medical payment fraud*: The emergence of tools that falsify medical management data to justify payments could enable Medicare and medical payment fraud.²⁹
5. *Scientific fraud*: The potential for scientific fraud may arise if scientific data are falsified, whether intentionally or not.³⁰
6. *Embedded prejudices*: Generative models are trained on large datasets, which may contain inherent biases or prejudices, leading to embedded racism and prejudices in the training data.³¹
7. *Data security challenges*: The storage and transfer of medical data are subject to data security issues, including unauthorized access or data breaches, which could have serious consequences, particularly if protected information is used as part of a training data set.
8. *Credentialing fraud*: As noted previously, generative models are almost capable of passing written credential examinations and could allow for proliferation of unqualified practitioners.⁷

Furthermore, models like ChatGPT are good at providing superficial summaries (as in the case of our summary), but often struggle with answering specific questions. For instance, when ChatGPT is prompted with “What are some of the perils of generative AI in neurointerventional surgery?”,

the generated answer is generic and identical to the one provided for a prompt where ‘neurointerventional surgery’ is replaced with ‘pediatrics’. These models lack introspection and do not have the ability to self-correct, nor do they have a fundamental understanding of the information on which they were trained, leading Bender *et al* to coin the term ‘stochastic parrots’ in order to describe large language models and arguing for careful curation of datasets rather than ingesting everything on the internet.³²

The question of high-quality datasets is an important point. The evolution, sophistication, and complexity of generative models has necessitated training on ever larger, high-quality datasets. GPT-3, Stable Diffusion, and DALL-E 2 models were trained on large publicly available datasets containing billions of data points. This presents serious ethical considerations regarding labor, content filtering, and sourcing as collecting and maintaining high-quality data is a daunting task, and annotating it is even more challenging.³³ Moreover, the scarcity of high-quality medical datasets underscores the critical importance of meticulous data adjudication, as models trained on subpar data can propagate errors when used to generate more extensive datasets. Some examples of neuroimaging datasets available are the OASIS-3 for Alzheimer's research and OpenNeuro, both of which are MRI only.^{34,35} CT datasets are even more rare, with one of the largest being the RSNA Brain Hemorrhage Detection Challenge dataset.³⁶ While large and well-annotated, this dataset does not include segmentation masks for hemorrhages, only label classification.

As an example of some of the issues with ensuring the quality of the data, the Brain Tumor Segmentation (BraTS) dataset is a large, well established database of brain MRI containing tumors.^{37–39} This dataset has been used in annual competitions in which various AI models are used to segment gliomas, and initially began with 65 clinical patients and 65 synthetically generated brain MRIs.³⁹ Subsequent additions to the BraTS dataset included ground truth masks created by the top performing algorithms of the BraTS competitions starting from 2013 onwards, and contributions from multi-institutional glioma collections of The Cancer Genome Atlas (TCGA), publicly available in The Cancer Imaging Archive (TCIA), in which ground truth masks were created and corrected by two imaging scientists and a medical physician. Incorrectly labeled voxels were corrected

following the rules set by an expert board-certified neuroradiologist, whose role was limited to approving the corrections.^{37,38} The implications of these approaches on generating data for machine learning are that datasets commonly used for training consist in part of synthetically generated data, with ground truth masks created by earlier generation top performing models from previous BraTS competitions, and in many cases were adjudicated by a single neuroradiologist who did not create or correct the ground truth masks. While generative AI has been proposed as a method for creating synthetic training data for machine learning, there have been several concerns about leakage of protected health information as well as a lack of metrics to ensure fidelity.^{40–43} Furthermore, using previous generation algorithms to create ground truths has an even higher concern for propagating bias, which ultimately may not be rigorously adjudicated by experts in the domain.

CONCLUSIONS

Generative AI is here to stay and will shape our future, whether we are ready for it or not. As generative models continue to be increasingly used in medicine, we must consider potential scenarios of abuse to prevent their occurrence. While these models have the potential to revolutionize medical practice, we must remain vigilant of the potential risks and consequences of their misuse. The possibility of a patient presenting deepfaked imaging and ChatGPT-generated medical reports in a clinical setting is no longer simply an academic thought exercise. The rapid pace of innovation in computer hardware and generative AI models continues to usher in increasingly impressive capabilities. Consumers can now generate high-quality images that were formerly prohibitively expensive to produce. Without adequate preparation, the consequences of artificially generated patient data would be significant to a surgeon's practice.

The field of medicine, particularly neurointerventional surgery, is continually evolving. While technology has significantly improved patient care, it took several decades and government mandates for electronic healthcare records to become widely used. The impact of ChatGPT and other generative algorithms on medicine is expected to be much more rapid.

According to Eric Topol, the greatest opportunity offered by AI is to restore the essential and time-honored connection and trust between patients and doctors.⁴⁴ We believe that there is

tremendous potential for this, particularly in the field of neurointerventional care. To ensure that this connection and trust is restored and protected, we must proactively marshal resources and regulations to effectively deploy or defend against generative algorithms. In the words of ChatGPT, "Fostering a collaborative and responsible approach between AI and healthcare providers could lead to significant advancements in patient care while preserving the human touch in medicine".

Twitter Tyler R Ray @raytusc and Ryan T Kellogg @RyanKelloggMD

Contributors TRR contributed to the conception of the work, drafting the work, and providing final approval of the version to be published. RTK contributed to drafting of the work and revising it critically for important intellectual content and providing final approval of the version to be published. KMF contributed to the drafting of the work and revising it critically for important intellectual content and providing final approval of the version to be published. FH contributed to the conception and drafting of the work and revising it critically for important intellectual content and providing final approval of the version to be published. JV contributed to the conception and drafting of the work and revising it critically for important intellectual content and providing final approval of the version to be published.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests RTK and JV have stock and are paid consultants for Viz.AI. KF serves on the editorial board of JNIS.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available. Not applicable.

© Author(s) (or their employer(s)) 2024. No commercial re-use. See rights and permissions. Published by BMJ.



To cite Ray TR, Kellogg RT, Fargen KM, et al. *J NeuroInterv Surg* 2024;**16**:4–7.

Accepted 26 June 2023
Published Online First 12 July 2023

J NeuroInterv Surg 2024;**16**:4–7.
doi:10.1136/jnis-2023-020353

ORCID iDs

Ryan T Kellogg <http://orcid.org/0000-0001-6491-5443>
Kyle M Fargen <http://orcid.org/0000-0001-8979-1993>
Ferdinand Hui <http://orcid.org/0000-0003-3759-7886>
Jan Vargas <http://orcid.org/0000-0001-7164-1479>

REFERENCES

- 1 Introducing ChatGPT. 2023 Available: <https://openai.com/blog/chatgpt>
- 2 Milmo D. ChatGPT reaches 100 million users two months after launch [The Guardian]. 2023. Available: <https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app>

- 3 Turing AM. Computing machinery and intelligence. *Mind* 1950;LIX:433–60.
- 4 Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks. *Commun ACM* 2020;63:139–44.
- 5 Karras T. Alias-free generative adversarial networks (StyleGAN3). 2023 Available: <https://nvlabs.github.io/stylegan3>
- 6 OpenAI. Improving language understanding with unsupervised learning. 2023. Available: <https://openai.com/research/language-unsupervised>
- 7 Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLOS Digit Health* 2023;2:e0000198.
- 8 Ali R, Tang OY, Connolly ID, et al. Performance of ChatGPT, GPT-4, and google bard on a neurosurgery oral boards preparation question bank. *Neurosurgery* 2023.
- 9 Cetinic E, She J. Understanding and creating art with AI: review and outlook. *ACM Trans Multimedia Comput Commun Appl* 2022;18:1–22.
- 10 Figurelle ME, Meyer DM, Perrinez ES, et al. Viz.ai implementation of stroke augmented intelligence and communications platform to improve indicators and outcomes for a comprehensive stroke center and network. *AJNR Am J Neuroradiol* 2023;44:47–53.
- 11 Murray NM, Unberath M, Hager GD, et al. Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: a systematic review. *J Neurointerv Surg* 2020;12:156–64.
- 12 Rava RA, Seymour SE, LaQue ME, et al. Assessment of an artificial intelligence algorithm for detection of intracranial hemorrhage. *World Neurosurg* 2021;150:e209–17.
- 13 Kellogg RT, Park MS, Snyder MH, et al. Establishment of age- and sex-specific reference cerebral ventricle volumes. *World Neurosurg* 2023;175:e976–83.
- 14 Kellogg RT, Vargas J, Barros G, et al. Segmentation of chronic subdural hematomas using 3D convolutional neural networks. *World Neurosurg* 2021;148:e58–65.
- 15 Lotan E. Emerging artificial intelligence imaging applications for stroke interventions. *AJNR Am J Neuroradiol* 2021;42:255–6.
- 16 Kather JN, Ghaffari Laleh N, Foersch S, et al. Medical domain knowledge in domain-agnostic generative AI. *NPJ Digit Med* 2022;5:90.
- 17 Nath S, Marie A, Ethershaw S, et al. New meaning for NLP: the trials and tribulations of natural language processing with GPT-3 in ophthalmology. *Br J Ophthalmol* 2022;106:889–92.
- 18 Zhang L, Dai H, Sang Y. Med-SRNet: GAN-based medical image super-resolution via high-resolution representation learning. *Comput Intell Neurosci* 2022;2022:1744969.
- 19 Ahmad W, Ali H, Shah Z, et al. A new generative adversarial network for medical images super resolution. *Sci Rep* 2022;12:9533.
- 20 Zhang K, Hu H, Philbrick K, et al. SOUP-GAN: super-resolution MRI using generative adversarial networks. 2021.
- 21 Tang J, LeBel A, Jain S, et al. Semantic reconstruction of continuous language from non-invasive brain recordings. *Nat Neurosci* 2023;26:858–66.
- 22 Parihar R, Dhiman A, Karmali T, et al. Everything is there in latent space: attribute editing and attribute style manipulation by styleGAN latent space exploration. MM '22; Lisbon, Portugal. New York, October 10, 2022:1828–36
- 23 Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: *Advances in neural information processing systems*. 2017. Available: <https://papers.nips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>
- 24 Rai A. Explainable AI: from black box to glass box. *J Acad Mark Sci* 2020;48:137–41.

- 25 Westerlund M. The emergence of deepfake technology: a review. *TIM Review* 2019;9:39–52.
- 26 Mirsky Y, Mahler T, Shelef I, et al. CT-GAN: malicious tampering of 3D medical imagery using deep learning. 2019.
- 27 Sousa Filho D de, Kanomata EY, Feldman RJ, et al. Munchausen syndrome and munchausen syndrome by proxy: a narrative review. *Einstein (Sao Paulo)* 2017;15:516–21.
- 28 Ekin T, Ieva F, Ruggeri F, et al. Statistical medical fraud assessment: exposition to an emerging field. *Inter Stat Rev* 2018;86:379–402.
- 29 Ranstam J, Buyse M, George SL, et al. Fraud in medical research: an international survey of biostatisticians. *Control Clin Trials* 2000;21:415–27.
- 30 Thorp HH. ChatGPT is fun, but not an author. *Science* 2023;379:313.
- 31 Korngiebel DM, Mooney SD. Considering the possibilities and pitfalls of generative pre-trained transformer 3 (GPT-3) in healthcare delivery. *NPJ Digit Med* 2021;4:93.
- 32 Bender EM, Gebru T, McMillan-Major A, et al. On the dangers of stochastic parrots: can language models be too big? In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. New York, NY, USA: Association for Computing Machinery, 2021: 610–23.
- 33 Exclusive: Openai used Kenyan workers on less than \$2 per hour to make Chatgpt less toxic; Available: <https://time.com/6247678/openai-chatgpt-kenya-workers> [Accessed 05 Mar 2023].
- 34 LaMontagne PJ, Benzinger TLS, Morris JC, et al. OASIS-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease. *Radiology and Imaging* 2019.
- 35 Markiewicz CJ, Gorgolewski KJ, Feingold F, et al. The OpenNeuro resource for sharing of neuroscience data. *Elife* 2021;10:e71774.
- 36 Flanders AE, Prevedello LM, Shih G, et al. Construction of a machine learning dataset through collaboration: the RSNA 2019 brain CT hemorrhage challenge. *Radiol Artif Intell* 2020;2:e209002.
- 37 Bakas S, Reyes M, Jakab A, et al. *Progression assessment, and overall survival prediction in the BraTS challenge*. 2019.
- 38 Bakas S, Akbari H, Sotiras A, et al. Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci Data* 2017;4:170117.
- 39 Menze BH, Jakab A, Bauer S, et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans Med Imaging* 2015;34:1993–2024.
- 40 Arora A, Arora A. Generative adversarial networks and synthetic patient data: current challenges and future perspectives. *Future Healthc J* 2022;9:190–3.
- 41 Chen RJ, Lu MY, Chen TY, et al. Synthetic data in machine learning for medicine and healthcare. *Nat Biomed Eng* 2021;5:493–7.
- 42 Frid-Adar M, Diamant I, Klang E, et al. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* 2018;321:321–31.
- 43 Guibas JT, Virdi TS, Li PS. Synthetic medical images from dual Generative adversarial networks. *ArXiv Preprint ArXiv:1709.01872* 2017.
- 44 Topol E. Deep medicine: how artificial intelligence can make healthcare human again. In: *Basic Books*. 2019. Available: <https://books.google.com/books?id=7iFIDwAAQBAJ>