limits of the theoretically derived estimates[8].

Jiang and co-workers' data, and data from similar studies, can aid estimates of the global effect of CO_2 fertilization. The size of this effect depends directly on the sensitivity of carboxylation efficiency to rising atmospheric CO_2 levels; this sensitivity should be similar at the eucalyptus forest and at all other sites around the globe, according to a theoretical analysis[8]. However, as CO_2 levels increase, the capacity of photosynthetic carboxylation to process more CO_2 diminishes, lowering the sensitivity of carboxylation efficiency to further CO_2-level increases. In other words, the CO_2-fertilization effect is dwindling at the biochemical level[8]. To work out the global fertilization effect, the carboxylation sensitivity is multiplied by the yearly increase in atmospheric CO_2 concentration, which is becoming larger over time. The yearly increase in CO_2 levels offsets the diminishing CO_2-fertilization effect.

Another factor that affects the size of the global CO_2-fertilization effect is the LAI[3]. The change in LAI observed at the Australian study site in response to CO_2 enrichment is at the low end of the wide spectrum of LAI changes that have been observed elsewhere[4–6]. At the global scale, however, the LAI is increasing over time — satellite observations show that Earth is literally becoming greener[9,10]. The increase of LAI amplifies the CO_2-fertilization effect.

The plant carbon-use efficiency reported in the current study is also at the low end of a wide range of reported values[7], and contributes to the low CO_2-fertilization effect observed in the study. However, we do not know much about how plant carbon-use efficiency varies over time at regional and global scales. This makes it difficult to assess whether the global fertilization effect will change because of shifts in this efficiency.

The bottom line is that it is currently difficult to estimate the size of the global CO_2-fertilization effect accurately. To solve this problem, we need to know more about hierarchical constraints not only across spatial scales, from ecosystem sites to regions and the globe, but also across biological scales — from the molecular level of biochemical reactions, to the leaf and canopy scale, and through to the larger scales associated with plant production and ecosystem carbon pools.

Yiqi Luo is at the Center for Ecosystem Science and Society, Department of Biological Sciences, Northern Arizona University, Flagstaff, Arizona 86011, USA. **Shuli Niu** is at the Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, and at the University of Chinese Academy of Sciences, Beijing 100101, China.
e-mail: yiqi.luo@nau.edu

1. Keeling, C. D., Chin, J. F. S. & Whorf, T. P. *Nature* **382**, 146–149 (1996).
2. Jiang, M. *et al. Nature* **580**, 227–231 (2020).
3. Li, Q. *et al. Biogeosciences* **15**, 6909–6925 (2018).
4. Duursma, R. A. *et al. Glob. Change Biol.* **22**, 1666–1676 (2016).
5. Ferris, R., Sabatti, M., Miglietta, F., Mills, R. F. & Taylor, G. *Plant Cell Environ.* **24**, 305–315 (2001).
6. Uddling, J., Teclaw, R. M., Kubiske, M. E., Pregitzer, K. S. & Ellsworth, D. S. *Tree Physiol.* **28**, 1231–1243 (2008).
7. DeLucia, E. H., Drake, J. E., Thomas, R. B. & Gonzales-Meler M. *Glob. Change Biol.* **13**, 1157–1167 (2007).
8. Luo, Y. & Mooney, H. A. in *Carbon Dioxide and Terrestrial Ecosystems* (eds Koch, G. W. & Mooney, H. A.) 381–397 (Academic, 1996).
9. Jiang, C. *et al. Glob. Change Biol.* **23**, 4133–4146 (2017).
10. Zhu, Z. *et al. Nature Clim. Change* **6**, 791–795 (2016).

Medical research

AI tracks a beating heart's function over time

Partho P. Sengupta & Donald A. Adjeroh

Clinicians use ultrasound videos of heartbeats to assess subtle changes in the heart's pumping function. A method that uses artificial intelligence might simplify these complex assessments when heartbeats are out of rhythm. **See p.252**

The heart is a specialized muscle that contracts rhythmically around its closed chambers to propel blood. However, this pumping function fluctuates throughout the day as the circulating blood flow adapts to the body's ever-changing metabolic demands[1]. Understanding the variations in cardiac pump activity with each heartbeat might have relevance for explaining the intricacies of heart function in health and disease. However, the tools for scrutinizing such changes remain imprecise. On page 252, Ouyang *et al.*[2] report the development of a computational platform that uses an artificial-intelligence (AI) approach to assess cardiac ultrasound video and to provide continuous, beat-by-beat measurement of cardiac pump function.

Clinicians commonly assess cardiac function using a value termed the ejection fraction, which is the percentage of the blood volume in the left heart chamber (the left ventricle) that is pumped out when the heart contracts. In a normal heart, just over half of the blood is ejected; thus, the calculated ejection fraction is more than 50%. Highly trained physicians can 'eyeball' ultrasound video loops of a beating heart and make a precise estimate of the ejection fraction[3]. However, if two isolated frames from the video were presented, showing only the beginning and the end of the ejection, even a trained physician would struggle to estimate the ejection fraction. Given that training and expertise vary from person to person, eyeballing is not relied on, and the ejection fraction is calculated by tracing the boundaries of the left ventricle on a digital image to estimate the blood volume at the beginning and end of ejection. It is recommended[4] that clinicians estimate the ejection fraction of a heart by tracking it over three to five heartbeats; however, in typical clinical practice, often just one beat is assessed.

If the accuracy of estimates of ejection fraction could be improved by having an easy way to routinely determine its precise value by tracking and averaging several heartbeats, this would be of immense benefit, particularly for people whose hearts are beating out of rhythm (a condition termed arrhythmia). If arrhythmia occurs, the changing duration of heartbeats alters the volume of blood filled and ejected from the left ventricle, thereby resulting in variations in the ejection fraction (Fig. 1). This variability makes the ejection fraction challenging to estimate for a type of arrhythmia known as atrial fibrillation. It is predicted[5] that this condition will affect between 6 million and 12 million people in the United States by 2050, and 17.9 million in Europe by 2060. Moreover, ejection fraction needs to be assessed frequently in people who have atrial fibrillation, because heart failure (a state characterized by a poor ability of the heart to pump blood) occurs in more than one-third of such individuals[6]. And more than half of people with heart failure have atrial fibrillation[6].

To develop an AI-based method for assessing ejection fraction, Ouyang *et al.* used 10,030 cardiac ultrasound videos. These videos were stored along with images containing human-generated tracings that marked the inner border of the left ventricle at the beginning and end of the ejection cycle. The authors used a type of AI architecture called a convolutional neural network (CNN), first to perform a semi-automatic detection of a pattern of pixel-based information (segmentation) to recognize the left ventricle in the video frames; and second, to track the borders of the ventricle during the heartbeat

cycle. Using CNN architecture to find the left-ventricle border in ultrasound images is not new[7,8], but the innovation here is that Ouyang and colleagues evaluated new forms of three-dimensional CNN. This enabled them to integrate recognition of the left-ventricular border (the spatial information) from the 2D display in single video frames with the changes over time (the temporal information), to determine the information needed regarding the moving heart border. Forms of 3D CNN have been used previously in realms as diverse as general video analysis[9,10], assessment of human physical activity[11] and medical imaging[12]. However, Ouyang and colleagues' work is, to our knowledge, the first attempt to take this approach in analysing cardiac ultrasound information over such a strikingly large number of videos.

After Ouyang and colleagues had 'trained' the 3D CNN using the video data, they compared the AI-generated estimates for the ejection fraction with human-measured ejection fractions. Their 3D CNN method estimated the ejection fraction with a mean error of 4.1% and 6%, respectively, for two different sets of data used for validation. In other words, on average, using the authors' proposed 3D CNN method, the ejection fraction was estimated to within 95.9% and 94%, respectively, of the corresponding ejection-fraction measurement reported by a clinician. These reported AI errors are substantially lower than those reported in previous attempts to use CNN to estimate the ejection fraction[7], and are well within the inter-observer variability in ejection-fraction measurements between experienced clinicians[3].

Ouyang *et al.* then tested a further 55 patients for whom 2 ultrasound specialists separately assessed the heartbeat videos. The authors found that when the variability in the human- and AI-generated ejection-fraction estimates for each patient was compared, the 3D CNN method produced results with the least variability in the ejection fractions noted between the two recorded measurements. Furthermore, results obtained with the 3D CNN were extremely consistent across different ultrasound machines, and for measurements taken on different occasions. These results also indicate the importance of assessing the kinetics of cardiac-wall motion in developing a system for gauging cardiac function.

Several avenues of possible future work building on this research should be explored. Efforts to reduce the overall computational burden would be welcome, so that the technique could be performed inexpensively and instantaneously during an ultrasound examination. Ouyang and colleagues' approach required 0.05 seconds per video frame, which they reported as being faster than the estimation speed of human experts. However, this is not yet as fast as real time, which would be less than 0.02 seconds per video frame (for a rate of 64 frames in 1.28 seconds). A careful look at the different stages in the overall architecture of the 3D-CNN deep-learning approach will be needed to determine the best architecture for use in existing cardiac-ultrasound technologies, such as 3D echocardiography and ultrafast cardiac ultrasound. Moreover, the choice of computational approach for handling videos that contain suboptimal images, or those in which the image quality has been improved by the injection of image-enhancing agents, will need to be considered.

This tool for the continuous assessment of cardiac pumping has the potential to affect other areas of cardiology. For example, such an approach might be adapted to monitor ultrasound changes in ejection fraction in people undergoing complex medical procedures, such as catheter-based cardiac interventions, surgery or when receiving medication or mechanical circulatory support for a condition termed acutely decompensated heart failure.

Furthermore, the use of 3D CNN to track other parameters that are more sensitive than ejection fraction for determining early changes in cardiac function (such as physical measures of heart-muscle deformation or changes in cardiac shape or geometry) that develop before a person shows disease symptoms might lead to new ways of measuring or identifying cardiac biomarkers (hallmarks of disease)[13–17]. Such automated approaches might be particularly relevant for the burgeoning 'multi-omics' approaches for data integration that incorporate different layers of biological information to define different stages of cardiac dysfunction[18].

In this regard, we applaud the authors for making available to the research community a large data set of annotated ultrasound videos (presented stripped of information that could identify the individuals). This resource will be extremely useful, and will probably spur yet more innovations in automated analysis that will boost our understanding of cardiac function. Moreover, such steps will be needed to achieve greater consistency in results obtained using different imaging systems for assessing cardiac function (such as cardiac ultrasound, computed tomography and magnetic resonance imaging).

The ongoing efforts to improve the accuracy of automated measurements and disease prediction will, undoubtedly, ultimately free up extra time for physicians, enabling them to provide higher-quality clinical care and have better interactions with patients. Given the high health-care burden of cardiovascular disease worldwide, Ouyang and colleagues' work is timely, and hints at an ensuing technological revolution that could have a profound effect on risk prediction of cardiovascular disease and on routine clinical decision-making.

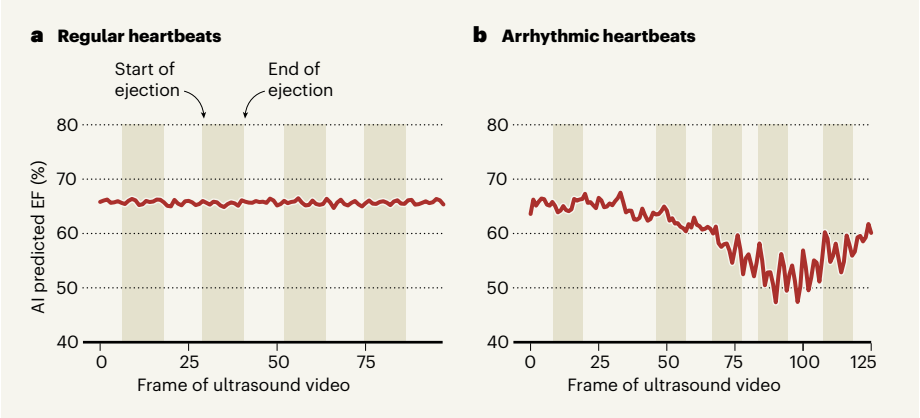

Figure 1 | A computational approach for assessing cardiac pump function over several heartbeats. Ouyang *et al.*[2] report the development of a method that uses artificial intelligence (AI) to monitor a standard clinical measurement of cardiac function. This measurement is called the ejection fraction (EF), and it is the percentage of the volume of blood in the left ventricle (one of the heart's chambers) that is pumped out of the chamber when the heart contracts. The authors developed a system that can analyse ultrasound video frames to determine the EF by continuously comparing changes in the borders of the left ventricle over space and time, using an AI architecture called a 3D convolutional neural network. **a**, Regular heartbeats have a fairly uniform EF. **b**, By contrast, irregularities (arrhythmic heartbeats) might result in shorter heartbeat cycles and a lower EF. The automation of heartbeat assessment by the use of AI (rather than depending on a clinician to monitor this) would make it easier to track several heartbeats. The averaging of EF over multiple heartbeats would provide a better heartbeat assessment for a person with arrhythmia than would be the case for the measurement of a single heartbeat. (Graphs based on Fig. 3 of ref. 2.)

Partho P. Sengupta is in the Department of Medicine, Division of Cardiology, West Virginia University Heart and Vascular Institute, West Virginia University School of Medicine, Morgantown, West Virginia 26505, USA. **Donald A. Adjeroh** is in the Lane Department of Computer Science and Electrical

Engineering, West Virginia University, Morgantown, West Virginia 26506, USA.
e-mail: partho.sengupta@wvumedicine.org

1. Martino, T. A. & Young, M. E. *J. Biol. Rhythms* **30**, 183–205 (2015).
2. Ouyang, D. *et al. Nature* **580**, 252–256 (2020).
3. Pellikka, P. A. *et al. JAMA Netw. Open* **1**, e181456 (2018).
4. Lang, R. M. *et al. Eur. Heart J. Cardiovasc. Imaging* **16**, 233–270 (2015).
5. Morillo, C. A., Banerjee, A., Perel, P., Wood, D. & Jouven, X. *J. Geriatr. Cardiol.* **14**, 195–203 (2017).
6. Santhanakrishnan, R. *et al. Circulation* **133**, 484–492 (2016).
7. Zhang, J. *et al. Circulation* **138**, 1623–1635 (2018).
8. Ghorbani, A. *et al. NPJ Digit. Med.* **3**, 10 (2020).
9. Ji, S., Xu, W., Yang, M. & Kai, Y. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**, 221–231 (2013).
10. Tran, D. *et al. Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recog.* 6450–6459 (2018).
11. Rahman, S. A. & Adjeroh, D. A. *Proc. IEEE Int. Conf. Bioinform. Biomed.* 1100–1103 (2019); https://doi.org/10.1109/BIBM47256.2019.8983251
12. Dolz, J., Desrosiers, C. & Ayed, I. B. *NeuroImage* **170**, 456–470 (2018).
13. Voigt, J.-U. *et al. J. Am. Soc. Echocardiogr.* **28**, 183–193 (2015).
14. Suinesiaputra, A., McCulloch, A. D., Nash, M. P., Pontre, B. & Young, A. A. *Med. Image Anal.* **33**, 38–43 (2016).
15. Varano, V. *et al. Med. Image Anal.* **46**, 35–56 (2018).
16. Bernardis, E. *et al. IEEE Trans. Biomed. Eng.* **65**, 733–744 (2018).
17. Omar, A. M. S. *et al. JACC Cardiovasc. Imaging* **10**, 1291–1303 (2017).
18. Cho, J. S. *et al. JACC Cardiovasc Imaging* https://doi.org/10.1016/j.jcmg.2020.02.008 (2020).

P.P.S. declares competing financial interests: see go.nature.com/38xpbrd for details.

This article was published online on 25 March 2020.

Tumour biology

Mutations in colon cancer match bacterial signature

Ye Yang & Christian Jobin

Studies have pointed to a link between colon cancer and a gut bacterium that produces DNA-damaging molecules. The discovery of a mutational signature linked to these bacteria in human colon cancer supports this association. **See p.269**

Understanding what causes colorectal cancer (CRC) could help to combat this disease of the colon. On page 269, Pleguezuelos-Manzano *et al.*[1] report evidence that strengthens a previously suspected connection to a type of gut bacterium. The authors implicate this microbe by pinpointing bacterial 'fingerprints' in DNA alterations found in CRC cells.

Certain bacteria produce genotoxic molecules — those capable of damaging DNA. These molecules can cause mutations if, for example, mistakes occur during the DNA-repair process that tries to fix genotoxic damage. In 2006, a genotoxin called colibactin, which is made by certain strains of the gut-dwelling bacterium *Escherichia coli*, was discovered[2]. That original description also shed light on how colibactin is produced by *E. coli*, identifying a key region of bacterial DNA, termed the *pks* island (the microbes that have this island are called *pks*$^+$ *E. coli*), which encodes various components of an 'assembly line' that makes colibactin.

By producing colibactin, *pks*$^+$ *E. coli* can accelerate tumour formation in animal models[3]. Moreover, these bacterial strains are more prevalent in close association with the epithelial cells in the mucosa of the colon in people who have CRC than in those who don't[3]. However, the complexity of the colibactin-producing assembly line and the molecule's considerable instability pose substantial challenges to researchers trying to decode the workings of the *pks* island and to characterize colibactin's structure.

There are several questions to be answered. For example, what is the mechanism of action of colibactin? What types of change might it make to DNA nucleotides? And does colibactin activity have relevance to human cancer? It is known[4] that *pks*$^+$ *E. coli* damages the DNA of cells it infects by causing adenine nucleotides to undergo a type of modification called alkylation. Subsequent evidence proposing a symmetrical colibactin structure indicates that the molecule has two 'warheads' made of a structure called cyclopropane, which target adenines[5]. How common *pks*$^+$ *E. coli* is in the gut of human populations is not fully known.

> **"The findings depict a potential pathway by which a genotoxic bacterium could contribute to the development of cancer."**

To determine the details of DNA changes that might be induced by *pks*$^+$ *E. coli*, Pleguezuelos-Manzano and colleagues turned to a 'mini-gut' cellular system that mimics the human intestine (Fig. 1). This approach uses a clump of human epithelial cells grown *in vitro* called an organoid or, specifically, a colonoid, because it is made of colon cells. The authors exposed colonoids either to *pks*$^+$ *E. coli* isolated from a person with CRC or to an engineered version of the bacterium that did not make colibactin. This set-up enabled the bacteria to interact with the type of cellular surface they would encounter in the lumen of the colon. Whole-genome sequencing of colonoid cells enabled the authors to compare the mutations in cells exposed to *E. coli* that produced colibactin or that were defective in producing it.

From this analysis, the authors determined a unique colibactin mutational signature — specific patterns of DNA alterations that arose in the presence of colibactin. This signature predominantly included two types of change. One type is the substitution of a single DNA nucleotide base for a different nucleotide (single-base substitutions, termed SBS-*pks*). These were skewed towards a change described as T→N, in which a thymine (T) nucleotide is replaced by any other type of nucleotide (N).

The other type of change is a small insertion or deletion of nucleotides, characterized by deletions of single nucleotides in stretches of thymine nucleotides (known as T homopolymers). This sort of alteration is termed ID-*pks*. Interestingly, both SBS-*pks* and ID-*pks* occur preferentially downstream of adenine nucleotides, consistent with the proposed mode of action of colibactin, with two warheads targeting adenine nucleotides that are located in close proximity on opposite strands of the DNA (one warhead targets an adenine upstream of the site of damage and the other targets the site of damage)[4,5].

To determine whether this colibactin-associated mutational signature might be relevant to human disease, the researchers analysed a data set[6] of whole-genome sequences for 496 human CRC tumours that had migrated from their primary site in the colon to form secondary growths termed metastases. Remarkably, the authors found that SBS-*pks* and ID-*pks* mutations were present in 7.5% and 8.8%, respectively, of CRC metastases, which is more frequent than in metastases of cancers of other primary origins. For example, SBS-*pks* and ID-*pks* mutations were found in 2.1% and 4.2%, respectively, of metastases of urinary-tract cancers, and in 1.6% and 1.6%, respectively, of head and neck tumour metastases. This pattern is consistent with the probability of exposure to *pks*$^+$ *E. coli* at these different body sites, considering that the urinary tract, head and neck are only occasionally exposed to *E. coli*. When the authors assessed 2,208 predominantly primary CRC tumours from an independent data set (see go.nature.com/3d6utsx), 5.0% and 4.4% of the tumours, respectively, had high SBS-*pks* and ID-*pks* signatures, which supports the idea that *pks*$^+$ *E. coli* are involved in