

Exploring the potential of generative artificial intelligence in medical image synthesis: opportunities, challenges, and future directions

Bardia Khosravi, Saptarshi Purkayastha, Bradley J Erickson, Hari M Trivedi, Judy W Gichoya



Generative artificial intelligence has emerged as a transformative force in medical imaging since 2022, enabling the creation of derivative synthetic datasets that closely resemble real-world data. This Viewpoint examines key aspects of synthetic data, focusing on its advancements, applications, and challenges in medical imaging. Various generative artificial intelligence image generation paradigms, such as physics-informed and statistical models, and their potential to augment and diversify medical research resources are explored. The promises of synthetic datasets, including increased diversity, privacy preservation, and multifunctionality, are also discussed, along with their ability to model complex biological phenomena. Next, specific applications using synthetic data such as enhancing medical education, augmenting rare disease datasets, improving radiology workflows, and enabling privacy-preserving multicentre collaborations are highlighted. The challenges and ethical considerations surrounding generative artificial intelligence, including patient privacy, data copying, and potential biases that could impede clinical translation, are also addressed. Finally, future directions for research and development in this rapidly evolving field are outlined, emphasising the need for robust evaluation frameworks and responsible utilisation of generative artificial intelligence in medical imaging.

Introduction

Generative artificial intelligence is a class of deep learning models capable of creating content that diverges from traditional discriminative models focused on interpretation or decision making. Generative artificial intelligence has seen rapid advancements over the past 3 years, with large language models gaining substantial public attention after the introduction of ChatGPT, a model trained on an extensive corpus of text to create coherent and realistic responses to user queries.¹ Large language models have shown noteworthy capabilities in the understanding and generation of natural language, paving the way for more advanced multimodal models that combine textual, visual, and contextual understanding. These large multimodal models have the potential to aid various domains, including health care, by integrating data from different input streams. Notable examples of large language models in medicine are Med-PaLM and Med-Gemini, which have shown promising results in tasks such as answering medical questions, summarising medical documents, and suggesting potential differential diagnoses on the basis of patient symptoms and test results. In addition, Med-Gemma and MedImageInsight are models trained on different types of medical images including radiology images (eg, chest x-rays, mammograms, CT), as well as dermatology and ophthalmology images, which allow end users to interact with the model using both language and images (and are thus known as multimodal foundation models). These multimodal models provide unconventional visual question answering ability and are able to learn from a few examples to perform downstream classification tasks.^{2,3}

Preliminary evidence suggests that generative artificial intelligence in the realm of visual content has made

remarkable advancements with models such as DALL-E, Stable Diffusion, Sora, and Veo, which excel in generating realistic images and videos based on textual prompts.⁴⁻⁶ Although these models primarily process text as input, with some using images for conditioning purposes, their primary focus is on generating high-quality images. Seminal works published since 2022 in medical imaging have shown the potential of generative artificial intelligence in creating realistic medical images (synthetic data), suggesting new approaches for research and clinical applications.⁷⁻¹⁰

This Viewpoint provides a comprehensive overview of synthetic data in medical imaging and critically analyses the advancements, applications, and challenges of this field. To this end, various image generation paradigms are examined, with the intention to assess how these generative technologies are changing the landscape of medical imaging research. The potential of these models and their derivative synthetic datasets, particularly their ability to augment and diversify medical research resources, are explored, in addition to their benefits in terms of data augmentation, anonymisation, and modelling biological phenomena. Finally, the challenges of using synthetic data are discussed, including the need for rigorous evaluation metrics and ethical considerations, and potential research directions are proposed that could substantially benefit the field of medical imaging.

Synthetic datasets

Generative models

The field of synthetic data is still in its nascent stages, with no consensus on a single, universally accepted definition as yet. This absence of a clear definition has led to

Lancet Digit Health 2025;

7: 100890

Published Online August 14, 2025
<https://doi.org/10.1016/j.landig.2025.100890>

Department of Radiology (B Khosravi MD MPH, Prof B J Erickson MD PhD) and Department of Orthopedic Surgery (B Khosravi), Mayo Clinic, Rochester, MN, USA; Luddy School of Informatics and Computing, Indiana University, Indianapolis, IN, USA (S Purkayastha PhD); Department of Radiology and Imaging Sciences, Emory University School of Medicine, Atlanta, GA, USA (H M Trivedi MD, J W Gichoya MD MS)

Correspondence to: Dr Judy W Gichoya, Department of Radiology and Imaging Sciences, Emory University School of Medicine, Atlanta, GA 30322, USA
judywawira@emory.edu

For Med-PaLM, see <https://sites.research.google/med-palm/>

For Med-Gemini, see <https://research.google/blog/advancing-medical-ai-with-med-gemini/>

For Med-Gemma, see <https://deepmind.google/models/gemma/medgemma/>

For MedImageInsight, see <https://learn.microsoft.com/en-us/azure/ai-foundry/how-to/healthcare-ai/deploy-medimageinsight>

For DALL-E, see <https://openai.com/index/dall-e-3/>

For Stable Diffusion, see <https://stability.ai/stable-image>

For Sora, see <https://openai.com/sora/>

For more on Veo, see <https://deepmind.google/technologies/veo/veo-2/>

inconsistencies in how the term is used and interpreted across various contexts, which in turn can affect the reproducibility and transparency of research involving synthetic data.¹¹ The Royal Society and The Alan Turing Institute put forth a working definition of synthetic data in 2022, as data that have been generated using a purpose-built mathematical model or algorithm, with the aim of solving a (set of) data science task(s).¹² This proposed definition emphasises the functional and intentional aspects of synthetic data, focusing on its strategic application in tackling complex scientific challenges rather than simply mimicking the statistical properties of the original data.

The advancement of generative artificial intelligence introduces a new concept in data sharing, which we refer to as a model as a dataset. In this concept, generative models learn and store patterns and characteristics of the original data in their internal parameters (weights).¹³ These trained weights contain a compressed version of the key features and relationships of the training data. Unlike traditional dataset sharing, which involves transferring actual images, sharing model weights provides an efficient alternative that allows others to generate new synthetic images with properties similar to the original data. These synthetic datasets have been shown to closely resemble the source data and capture their distribution, including the relationship of different anatomical features and their correlation with different pathological processes.^{8,9}

Two broad categories of generative models provide the ability to generate synthetic datasets: physics-informed and statistical models.

Physics-informed models are primarily rule-based approaches that incorporate domain-specific knowledge and physics principles through mathematical equations and explicit constraints to generate realistic and physically plausible data. Rather than learning the patterns directly from data, these models encode expert knowledge and known physics laws (eg, fluid dynamics, tissue biomechanics, or radiation physics) to simulate biological phenomena. These models have been applied successfully in medical imaging to simulate anatomical structures (such as a shape model of the femoral bone), physiological processes (such as blood flow dynamics in vascular structures), and medical interventions (such as simulating the distribution of the radiation dose in radiotherapy planning).¹⁴ Physics-informed models offer high fidelity and interpretability but might require extensive domain expertise and computational resources.

In contrast to physics-informed models, statistical models learn from data patterns and distributions (figure 1). Among them, variational autoencoders (VAEs) function by compressing data into a lower-dimensional representation, also known as latent space, and then reconstructing the data, thereby capturing the data distribution effectively.¹⁵ Generative adversarial networks (GANs) operate through a dual-network system, in which a generator creates data samples and a discriminator evaluates these data samples and provides feedback to the generator.¹⁶ This synergy

continually enhances the quality and realism of the data generated. Denoising diffusion probabilistic models (DDPMs) introduce noise into an image and learn to reverse this process, producing high-quality samples.¹⁷

Statistical models encounter the generative artificial intelligence trilemma, which involves balancing high sample quality, comprehensive mode coverage, and rapid sampling rates (figure 2).¹⁸ VAEs are notable for their quick sampling capabilities, sometimes resulting in lower sample quality. GANs excel at generating high-quality samples but might not always capture all data variations, leading to low mode coverage, known as mode collapse. DDPMs stand out for their ability to generate samples of exceptional quality and extensive mode coverage, albeit at a slower sampling rate. End users select the generative model that matches their application of interest, balancing the desired image quality and speed. For dataset generation purposes, the priority typically shifts towards ensuring high image quality and comprehensive mode coverage, often outweighing concerns of sampling speed.

Use cases in medical imaging

Generative models and their synthetic datasets have numerous applications in medical imaging (panel 1). One well studied use case involves supplementing or replacing real data to train deep learning models for downstream tasks such as classification or segmentation. Generated images can be conditioned on class labels (eg, presence or absence of pneumonia) or descriptive text (eg, right middle lobe consolidation). Research has shown that images generated by GANs and DDPMs can improve the performance of downstream pathology classifiers substantially.^{7,19,30} Notably, the classifier performance improves as more synthetic data are added to the real dataset. In some cases, a sufficiently large pool of generated images can match the performance benefit of real data, potentially opening new avenues for data sharing whereby synthetic data acts as a replacement of the original data.⁸ However, when training and evaluating generative models, caution is required to avoid distribution leakage (in which a patient is represented in both training and test data), which could overestimate performance improvements.⁸ Of note, repeatedly training image generation models on the output of other generative models (usually more than three iterations) risks mode collapse, which degrades the quality of the final model.³¹

Generative models also excel at image transformations. VAEs and GANs have long enabled low-dose CT image denoising, eventually reducing radiation exposure for patients.^{32,33} Of late, accelerated MRI techniques have been used to reduce the scan time by 30%.³⁴ Another image-to-image transformation use case generates missing MRI sequences, enabling training of downstream algorithms requiring all four sequences: T1, T2, post-contrast T1, and FLAIR.^{23,29} DDPMs have enabled inpainting, which involves selectively adding or removing specific image parts on the basis of criteria, without altering the context. For instance, trained diffusion models can introduce brain tumour

lesions in healthy brain MRIs or remove tumoural regions by drawing on an image.²⁴ Such edits can enrich under-represented datasets and introduce rare conditions, such as adding brain tumours to individuals with Alzheimer's disease. A more advanced version of the inpainting technique was developed to edit specific regions of a chest radiograph using text prompts.²⁶ The resulting edited images were used to stress-test existing models—for example, removing chest tubes from pneumothorax images to evaluate classifier performance without this known confounder.³⁵

Evaluating image quality

Evaluating the quality of generated images, which determines how these synthetic images are used, is crucial. Various metrics have been proposed to quantify the quality of generated images, both in the presence and absence of ground truth references. These metrics can be broadly categorised into two groups: image metrics and text-image metrics (panel 2).

Image metrics

When ground truth images are available—for example, in tasks such as super resolution and denoising of medical images—traditional metrics such as structural similarity index and the peak signal-to-noise ratio can be used to measure the similarity between the generated and reference images.^{36,37} However, in the absence of ground truth—for example, in class-conditioned image generation—alternative metrics are required. For instance, classification accuracy score trains a classification model on derived medical data and evaluates its performance on real images, providing insights into the domain adaptation capabilities of the generation models.³⁸

Another widely adopted metric is the inception score, which uses an inception network pretrained on ImageNet to evaluate class predictions for a set of generated samples.³⁹ Fréchet inception distance (FID) compares the means and covariances of features extracted by an ImageNet-pretrained inception network between the generated and real samples.⁴⁰ By accounting for the target distribution, FID provides a better estimate of image diversity than inception score. Several variants and improvements of FID have been proposed—eg, the kernel inception distance is a variant of FID that enables metric calculation using a small number of samples, unlike FID calculation, which requires generation of a large number of samples and is resource intensive.⁴¹ One limitation of these metrics is that they depend on pre-trained networks, and unlike natural images, no universally accepted model for feature extraction exists in medical imaging.

Human evaluation

In addition to computational metrics, human evaluation remains a gold standard for assessing the quality of generated medical images. The human Turing test involves domain experts who are asked to discern between real and

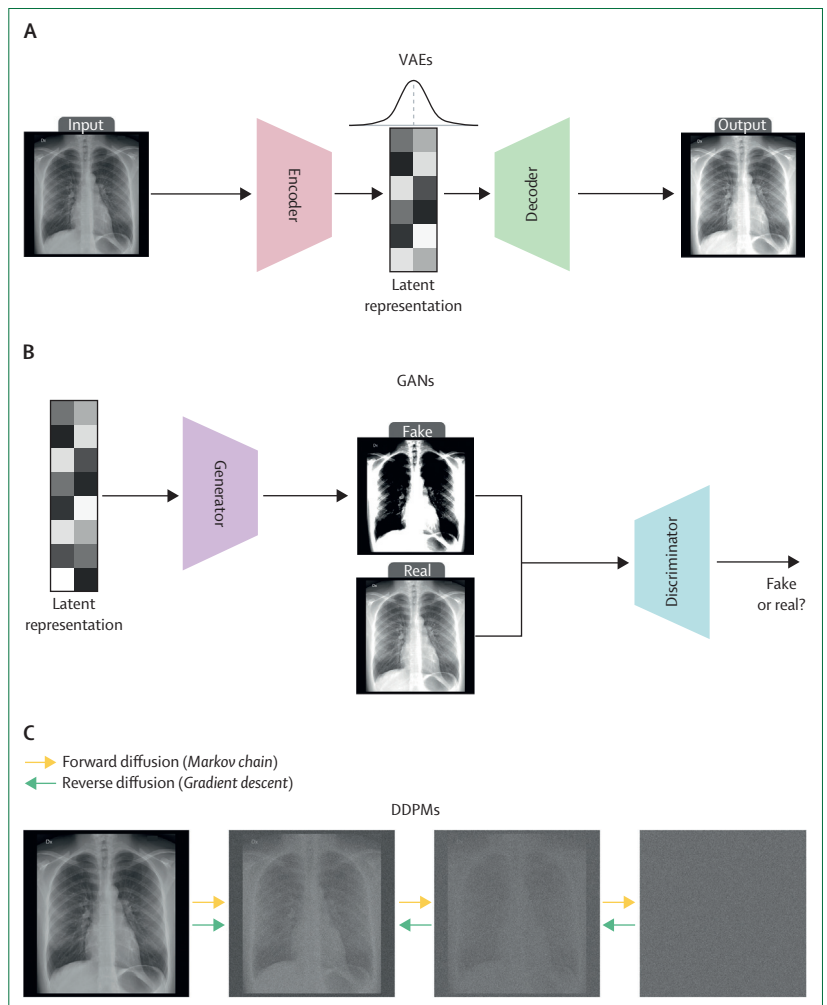


Figure 1: The architectures and key components of three popular statistical models used in image generation (A) VAEs consist of an encoder that compresses the input data into a lower-dimensional latent representation and a decoder that reconstructs the original data from the latent space. The model is trained to minimise the reconstruction error while also regularising the latent space to follow a previous distribution, typically a standard normal distribution. This training enables the generation of new samples by sampling from the learned latent distribution and decoding them. (B) GANs use a two-network architecture, with a generator that creates synthetic data samples and a discriminator that distinguishes between real and generated samples. The generator aims to produce samples that are indistinguishable from real data, whereas the discriminator provides feedback to guide the generator's improvement. Through an adversarial training process, the generator learns to capture the underlying data distribution, enabling the creation of realistic samples. (C) DDPMs generate data by learning to reverse a noising process. The model starts with a sample from a simple distribution (eg, Gaussian noise) and iteratively denoises the sample using a learned Markov chain. At each step, the model estimates the gradient of the data distribution and refines the sample accordingly. By repeatedly applying this process, DDPMs can produce high-quality samples that closely resemble the training data. The figure depicts the forward diffusion process that gradually adds noise to the data and the reverse diffusion process that progressively denoises the sample to generate a clean output. DDPMs=denoising diffusion probabilistic models. GANs=generative adversarial networks. VAEs=variational autoencoders.

derived medical images.⁴² This assessment provides insights into the perceptual quality and realism of generated images, which is crucial for medical imaging, in which accuracy and fidelity are paramount. However, as perceptual quality and realism are subjective measures, a wide range of participants with different experience levels should be involved in the image evaluation process.⁴³

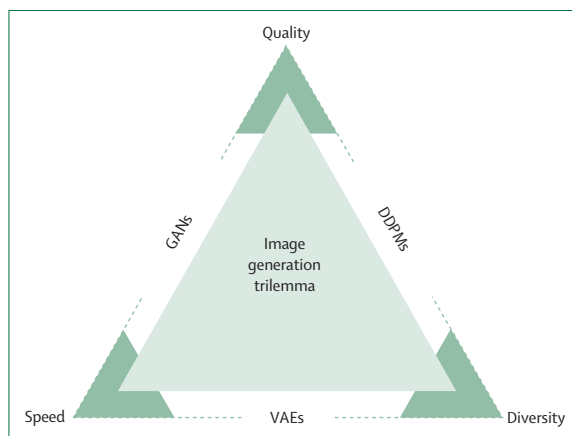


Figure 2: The image generation trilemma, which represents the trade-offs between three key aspects of generative models: diversity, quality, and speed VAEs excel in generating diverse samples quickly but can compromise on image quality. GANs strike a balance, providing good quality and diversity but can suffer from mode collapse, thereby restricting the diversity. DDPMs prioritise high-quality and diverse samples at the cost of a slow generation speed. DDPMs=denoising diffusion probabilistic models. GANs=generative adversarial networks. VAEs=variational autoencoders.

Text-image metrics

Although image metrics focus solely on the visual quality of generated images, text-image metrics aim to measure the alignment between the input text and the generated image. These metrics are particularly relevant in medical image generation tasks, in which the generated images need to reflect the textual descriptions of medical conditions or anatomical structures accurately. Metrics such as contrastive language-image pretraining score (CLIPScore) and bootstrapping language-image pretraining score (BLIP-Score) measure the similarity between the input text and the generated image, quantifying the degree of alignment between the two modalities.^{44,45}

Image-text matching is another crucial group of metrics for evaluating the alignment between generated medical images and their corresponding textual descriptions. Compositional quality metrics assess this alignment by decomposing the text and image into individual components and measuring their correspondence, often using object detection techniques.^{46,47} These metrics go beyond overall visual similarity and focus on accurate representation of specific anatomical structures, pathologies, or medical conditions mentioned in the text. By ensuring that the generated images convey the intended medical information accurately, compositional quality metrics can play a key role in medical education and research.

Health-care-specific metrics

Evaluating synthetic medical images requires metrics tailored to health-care needs, beyond general purpose tools such as the structural similarity index or FID. Efforts are under way to adapt existing metrics for medical contexts. For instance, researchers have begun replacing ImageNet-pretrained models in FID with networks trained on medical

datasets such as RadImageNet to create a medical FID, which captures the statistical properties of radiology images better.⁴⁸ However, health-care-specific metrics remain an active research area as disease classifiers might rely more on local features than global features.⁴⁹ Similarly, anatomical accuracy is being prioritised by developing measures that use segmentation tools to ensure that crucial structures (such as organs or lesions) are preserved in synthetic images.⁵⁰ These adaptations aim to address the limitations of standard metrics, which often fail to reflect clinical relevance or diagnostic utility.

A suggested next step is to integrate clinical validation with these computational approaches. Human evaluations such as the human Turing test already involve experts distinguishing real images from synthetic ones, offering insights into the perceptual quality that is important for medical use. For text-guided image generation, metrics such as CLIP-Score are being refined by using medical foundation models such as BioMedClip.⁵¹ Testing synthetic images in practical, clinical tasks such as training classifiers for disease detection can further highlight their utility. Combining these efforts could provide a robust, health-care-specific evaluation, thereby ensuring that synthetic images meet both technical and clinical standards for advancing medical imaging research and practice.

Potentials and promises

Synthetic data generation and image generation models hold immense promise for the future of medical imaging research. By leveraging the power of generative models, researchers can unlock unprecedented levels of data diversity, privacy preservation, and multifunctionality, changing the way dataset creation, utilisation, and disease modelling are approached.

Increased dataset size and diversity

One of the key advantages of generating data via statistical models lies in their ability to increase dataset size and diversity. Preliminary evidence suggests that generative models can be trained to disentangle specific associations within data, allowing for the creation of novel combinations that might not be readily available in real-world datasets.^{52,53} For instance, a model trained on brain MRI scans can generate images with varying degrees of atrophy or lesion load, independent of factors such as age or sex. Such disentanglement enables training models to detect specific pathologies without confounding the effects of other variables. As mentioned earlier, supplementing increased dataset size with generated images could lead to enhanced downstream model performance.⁸ Moreover, targeted oversampling of minoritised sociodemographic groups or patients diagnosed with rare diseases through synthetic data generation has been shown to close the fairness gap by 40%.²² Synthetic data generation closes this fairness gap by facilitating an increase in dataset sizes that represent the original dataset distribution for various subgroups.

Panel 1: Use cases of synthetic imaging datasets and image generation models in medicine and their findings

We performed a literature search using the following term combinations: "synthetic data" OR "VAE" OR "GAN" OR "diffusion model" AND "medical imag" OR "radiolog" OR "dermatolog" OR "patholog", and selected a representative paper matching each type of synthetic data for inclusion.

Chambon et al (2022):⁷ Generating chest radiographs conditioned on input prompts

- Improved classifier performance by 5% when trained on combined synthetic and real data
- Increased classifier performance by 3% when trained solely on a larger synthetic dataset
- Enhanced text encoder representation for pneumothorax detection by 25% after fine-tuning

Pinaya et al (2022):⁹ Generating 3D brain MRIs and investigating the conditioning factors

- Enabled controlled generation of realistic 3D brain MRIs with adjustable age, sex, and structural parameters
- Generated a synthetic dataset of 100 000 brain images for public use

Frid-Adar et al (2018):¹⁹ Generating abnormal samples to tackle class imbalance in liver lesion detection on CT scans

- Increased liver lesion detection sensitivity from 78.6% to 85.7% using synthetic data augmentation
- Improved specificity from 88.4% to 92.4% with synthetic images
- Radiologists found synthetic images indistinguishable from real ones in blinded assessments

Khosravi et al (2024):⁸ Using synthetic chest radiographs to supplement real images to expand the training set of pathology classifiers

- Enhanced AUROC by up to 0.02 in internal and external test sets with ten times synthetic data supplementation
- Synthetic-trained classifiers matched the performance of real-data models using 33–50% fewer images
- Combining real and synthetic data improved AUROC of pathology classifiers from 0.76 to 0.80 in cross-source testing

Ktena et al (2024):²⁰ Using synthetic images to increase the fairness of downstream classifiers on multiple modalities

- Reduced fairness gap by 44.6% in chest radiograph classifiers trained on synthetic and real images
- Improved out-of-distribution prediction accuracy by 7.7% across pathology slides
- Increased dermoscopy high-risk sensitivity by 63.5% and reduced fairness gaps by 7.5 times

Conte et al (2021):²¹ Creating missing brain MRI sequences for streamlined processing

- Boosted tumour segmentation Dice coefficient from 0.79 to 0.83 with synthetic MRI sequences

Rouzrokh et al (2022):²² Introducing and removing lesions from brain MRI slices

- Effectively inpainted (which involves selectively adding or removing specific image parts on the basis of criteria, without altering the context) tumour components, random tumours, and healthy brain tissues using DDPMs

Khosravi et al (2024):²³ Creating counterfactual pelvis radiographs from different race groups to evaluate disparities in large imaging datasets

- Identified racial disparities in prevalence of osteoarthritis between African American patients and White patients
- Highlighted dataset-scale disparities by means of synthetic counterfactual pelvis radiographs

Pérez-García et al (2023):²⁴ Stress-testing image classifiers by creating counterfactuals to evaluate possible shortcuts and their effect on model performance

- Generated counterfactual datasets simulating acquisition, manifestation, and population shifts
- COVID-19 classifier accuracy dropped from 99.1% to 5.5% when COVID-19 features were removed
- Pneumothorax classifier accuracy dropped from 93.3% to 17.9% when chest tubes were artificially removed

Khosravi et al (2023):²⁵ Using internal features of generative models for label-efficient pelvis radiograph segmentation

- Enhanced pelvis radiograph segmentation accuracy by 0.30–0.32 points using generative model features, using only 20 annotated samples

Rouzrokh et al (2024):²⁶ Creating synthetic postoperative images of patients undergoing total hip arthroplasty

- Produced synthetic postoperative hip radiographs with a mean acetabular angle of 39.9° (±4.6), 99% within safe zones
- Synthetic radiographs scored higher validity (9.0±0.7) than real ones (7.9±1.1)

Yuan et al (2024):²⁷ Imputing missing 3D brain MRIs in Alzheimer's longitudinal studies conditioned on past or future scans

- Achieved SSIM of 0.895 (with skull) and 0.983 (skull removed), outperforming autoencoders (0.74 for with skull and 0.91 for skull removed) and naive methods (0.70 for with skull and 0.89 for skull removed)
- Reduced volumetric error rates from 0.14 (using conventional methods) to 0.05

(Continues on next page)

Panel 1 (continued from previous page)

Kyung et al (2024):²⁸ Forecasting chest radiograph morphology on the basis of electronic health record data

- Achieved a weighted macro AUROC of 0.72 in predicting future chest x-ray pathologies, outperforming tabular-only classifiers and previous label baselines
- Maintained sex (AUROC 0.96) and age (0.45) correlations in synthetic images

Liu et al (2025):²⁹ Forecasting tumour growth on the basis of baseline tumour characteristics and treatment plan

- SSIM of 0.92 and PSNR of 29.0 for multiparametric MRI generation, outperforming baseline models without treatment-aware conditioning
- Generated MRI quality remained high across different treatment-day ranges, with SSIM ranging from 0.88 to 0.94 depending on the treatment phase
- Tumour growth predictions were most reliable within a 4-month window, with the Dice similarity coefficient dropping from 0.85 to 0.46 as the time interval extended from 0.5 months to greater than 24 months

3D=three dimensional. AUROC=area under the receiver operating characteristic curve. DDPMs=denoising diffusion probabilistic models. PSNR=peak signal-to-noise ratio. SSIM=structural similarity index.

Panel 2: Summary of image quality metrics based on use case for medical image generation

Image super resolution, denoising, and inpainting

- SSIM: assesses structural similarity between generated and reference images by considering luminance, contrast, and structure
- PSNR: measures the ratio between the maximum possible power of a signal and the power of corrupting noise between generated and reference images

Class-conditioned and unconditional image generation

- IS: compares class predictions and diversity of generated samples using an inception network pretrained on ImageNet
- FID: compares means and covariances of features extracted from generated and target distributions using an inception network pretrained on ImageNet
- KID: computes squared MMD between inception representations of generated and target distributions using an inception network pretrained on ImageNet

Domain adaptation, and class-conditioned image generation

- CAS: uses a classifier trained on derived medical images and evaluates performance on real images

Perceptual quality assessment, and realism evaluation

- Human Turing test: medical experts discern between real and derived images

Image generation from textual descriptions

- Segmentation-based metrics: volumetric analysis of different organs on generated images and comparing them with the input condition
- CLIPScore: computes cosine similarity between CLIP embeddings of text descriptions and generated images
- BLIPScore: computes cosine similarity between BLIP embeddings of text descriptions and generated images
- LLMscore: leverages an LLM for creating a detailed caption at the level of an image and different objects and compares the generated caption with the input text descriptions

BLIP=bootstrapping language-image pretraining. CAS=classification accuracy score. CLIP=contrastive language-image pretraining. FID=Fréchet inception distance. IS=inception score. KID=kernel inception distance. LLM=large language model. MMD=maximum mean discrepancy. PSNR=peak signal-to-noise ratio. SSIM=structural similarity index.

Privacy preservation

Synthetic datasets offer a privacy-preserving solution to the challenges of sharing and utilisation of data in medical research.⁵⁴ Generative artificial intelligence anonymises sensitive patient information by generating realistic images that mimic biological characteristics of real patient data (both visually and in the model feature space) without direct replication of original data.⁵⁵ Such anonymisation enables the creation of datasets that can be shared and analysed without compromising patient privacy, which further opens up new avenues for collaborative research and facilitates the development of robust, privacy-compliant artificial intelligence models in medical imaging.

Versatility across tasks

Another key potential of image generation models, especially DDPMs, lies in their multifunctional nature.

Generative models trained on medical images can be adapted and repurposed for various tasks beyond supplementing data; for example, features learned from an unsupervised image generation model can be leveraged for few-shot image segmentation, enabling accurate delineation of anatomical structures or pathologies with only 20 expert-annotated examples.²⁷ The same model without any further training can also be used for inpainting to create diverse training samples.⁵⁶ Similarly, generative models without any fine-tuning after initial training can be used for anomaly detection in medical images.^{57,58} This versatility extends the value of the synthetic datasets and their generator models, as a single model can be used for multiple downstream applications, streamlining research workflows, and reducing the need for task-specific data collection and model development.

Panel 3: Summary of challenges, considerations, and future research directions for generative artificial intelligence and synthetic datasets in medical imaging**Data copying**

Generative models can inadvertently reveal sensitive patient information when they reproduce images that closely resemble the original data.

Future research directions:

- Creating metrics to quantify the privacy risk of generated images
- Developing post-hoc data anonymisation methods
- Investigating the trade-off between image quality and privacy preservation

Identification of source dataset

Identifying specific datasets used to train generative models can be challenging, hindering the assessment of potential biases or limitations in the generated data.

Future research directions:

- Creating standardised reporting guidelines for synthetic medical imaging datasets
- Developing techniques for dataset fingerprinting in generative models
- Creating trusted third-party validation services for synthetic medical datasets
- Exploring methods for reverse-engineering model-training data

Interpretability and explainability

The complex nature of generative models makes understanding how these models learn and generate data a challenge. This understanding is necessary to build trust in the model outputs.

Future research directions:

- Implementing uncertainty quantification methods for stochastic prediction models
- Creating clinically relevant interpretability metrics
- Developing interactive visualisation tools for clinicians to explore model decisions
- Investigating the integration of domain knowledge into model explanations

Potential biases

Biases in the source datasets could be propagated or amplified in the generated data, leading to skewed research findings or discriminatory applications.

Future research directions:

- Creating benchmarks for evaluating fairness in medical imaging generative models
- Establishing multi-institution collaboratives to create demographically balanced training data
- Investigating the effect of data augmentation on bias reduction

Modelling complex biological phenomena

Advanced generative models can internalise complex biological phenomena through their training procedures, enabling the intricate physiological processes to be modelled and simulated.⁵⁹ This internalised world model can be leveraged for novel applications that extend beyond the downstream tasks discussed in the previous section. One striking example of this capability is the prediction of postoperative imaging appearances. When trained on a large corpus of paired prearthroplasty and postarthroplasty pelvic radiographs, these models generated highly realistic postoperative radiographs, simulating a well executed surgery.²⁸ Remarkably, domain-expert surgeons evaluated the generated postoperative images as more robust and anatomically accurate than their real counterparts, highlighting the potential of these models in serving as virtual surgical planning tools and educational resources.²⁸

Another compelling application of this internalised world model is the prediction of disease progression.³⁰ For instance, when given an initial brain MRI scan and information about the patient's treatment regimen, advanced DDPMs can generate a series of images that depict the potential progression of a brain tumour over time.³¹ By

learning the complex interplay between disease characteristics, treatment effects, and biological processes, these models can provide valuable insights into patient prognosis and aid clinical decision making.³¹

Challenges and considerations

Although derivative synthetic datasets and image generation models hold immense promise for medical imaging research, several challenges and ethical considerations need to be addressed to ensure their responsible and effective utilisation. Panel 3 summarises these challenges and proposes some future research directions to mitigate them.

Patient privacy and data copying

Although synthetic datasets can help to preserve patient privacy by generating anonymised data, concerns regarding potential data copying still exist.⁶⁰ If a generative model is trained on a specific dataset and can replicate images that closely resemble the original data, then the model might inadvertently reveal sensitive patient information. Copying happens when multiple copies of the image or captions are present in the dataset, which not only necessitates careful data curation,⁶¹ but also raises concern about the degree of

anonymisation achieved in the training data and the potential for reidentification. Unlike tabular data, medical images contain patient-identifying information embedded within the pixel values, thus posing unique challenges for anonymisation. For instance, facial features in brain MRIs or distinctive anatomical markers in radiographs might enable reidentification even when explicit patient identifiers are removed.^{62,63}

Researchers need to carefully assess the risk of data copying and implement measures to mitigate this concern, such as using differential privacy techniques or post-hoc data anonymisation.^{64,65} Advances made over the past 4 years in privacy evaluation metrics for synthetic data, such as membership inference attacks and similarity scores between real and generated samples, can help to quantify privacy risks. Additionally, emerging standards for synthetic content provenance, including the Coalition for Content Provenance and Authenticity (C2PA) and Google's SynthID, have been developed to label artificial intelligence-generated content, addressing both transparency and intellectual property concerns.⁶⁶

For more on C2PA, see <https://c2pa.org/>

For SynthID, see <https://deepmind.google/science/synthid/>

Identification of source dataset and disclosure

Transparency regarding the source datasets used to train generative models is crucial in ensuring the integrity and reproducibility of research findings. However, identifying the specific training datasets can be challenging, especially when models are trained on multiple proprietary sources or when researchers use pretrained models without full knowledge of their training data.⁶⁷ This insufficient transparency can hinder the ability to assess potential biases or limitations in the generated data. To address this gap, researchers should strive to document and disclose all source datasets used in the training process, enabling better understanding and validation of the derived data. Additionally, specific hyperparameters used for inference, specific class or prompt conditions, and every post-processing step involved in creating the synthetic dataset should be released along with the model or dataset release, to ensure reproducibility and applicability of the downstream work.⁶⁸ Dataset documentation guidelines, such as the STANDING Together guidelines published in 2024, should be adopted for synthetic data generation models.⁶⁹

Interpretability and explainability

As generative models become increasingly complex, their interpretability and explainability will become more challenging. Understanding how these models learn and generate data is crucial for building trust in their outputs and ensuring their safe and reliable use in medical imaging research. Although some specific explainability methods devised for generative models exist to ensure proper understanding of input text or to add uncertainty measures to the datasets, adaptation and evaluation of these methods in medical imaging remains restricted.^{70,71}

Potential biases

The use of synthetic datasets and generative models raises important bias considerations. The potential for biases in the source datasets getting propagated or amplified in the generated data is a key concern.⁷² If the training data are biased towards some demographics, pathologies, or imaging protocols, then the resulting generated data could perpetuate these biases, leading to skewed research findings or discriminatory applications.²⁵ For instance, historically, many medical imaging datasets have under-represented minoritised populations, resulting in artificial intelligence systems with likely differential performance levels across demographic groups.⁷³ When representation is low, generative models could struggle to capture a true distribution of these under-represented groups. However, a 2023 study suggests that newer generative models can arrive at meaningful representations from as few as 20 samples when the overall dataset is sufficiently large to capture high-level features.⁶¹ Mitigation strategies in this case include diversity-aware sampling during training, adversarial debiasing techniques, explicit fairness constraints in model objectives, and leveraging the few-shot fine-tuning capabilities of newer generative models.⁷⁴ Researchers need to actively assess and mitigate potential biases in the source data and regularly audit the generated data for fairness and representativeness.

Future directions

The field of generative artificial intelligence in medical imaging is evolving rapidly, and several key areas of research and development hold promise for advancing the capabilities and applications of synthetic datasets and image generation models. One crucial direction is the development of more robust and standardised evaluation frameworks that consider the unique challenges and requirements of medical imaging, including establishment of clinically relevant metrics, benchmark datasets, and challenges concerning comparative analysis and validation of different generative models.⁷⁵

Another important avenue is the exploration of novel architectures and training strategies, such as hybrid models combining physics-informed and statistical approaches with incorporation of domain-specific knowledge and constraints. Integration of generative models with other artificial intelligence techniques such as reinforcement and active learning could enable the creation of personalised and patient-specific datasets for precision medicine and targeted treatment planning.⁷⁶

Addressing the ethical and regulatory challenges surrounding the use of synthetic datasets and image generation models is essential to realise their full potential, and requires collaboration among researchers, clinicians, ethicists, and policy makers to develop guidelines and best practices for responsible use, data privacy, consent, and accountability. Regulatory bodies, including the US Food and Drug Administration (FDA) and the European Medicines Agency, will play a crucial role in establishing

frameworks for validating and approving synthetic data for clinical applications. Frameworks for evaluating synthetic medical imaging are already emerging, as evidenced by the FDA's clearance of synthetic MRI technologies.⁷⁷ These technologies were regulated as image processing software rather than as completely novel modalities, with the FDA requiring extensive clinical validation to show that the diagnostic performance of the radiologist remained equivalent when using synthetic images versus conventional images. This regulatory precedent suggests a pathway for future synthetic data technologies: proof-of-performance equivalence on standardised diagnostic tasks, rigorous clinical validation with multiple readers, and postmarket surveillance commitments to monitor for any divergence in clinical outcomes.

In conclusion, derivative synthetic datasets and image generation models have the potential to change medical imaging research and clinical practice. Addressing the challenges associated with them, establishing best practices, and investing in research and innovation can help to harness the full potential of generative artificial intelligence in improving patient care, advancing scientific discovery, and transforming the landscape of medical imaging.

Contributors

BK and JWG conceptualised the Viewpoint and wrote the original draft. All authors contributed to the review and editing of the manuscript. All authors have read and approved the final version of the manuscript.

Declaration of interests

BK and BJE have pending patents on radiographic image generation and feature extraction from generative models (63/583044 and PCT/US2023/074166). JWG is a member of the American College of Radiology Artificial Intelligence Advisory Group, the Society of Imaging Informatics in Medicine Board, and the Health Level 7 Standards Board; has received honoraria from the National Bureau of Economic Research for writing in their 2023 conference book; and has grants or contracts from Lunit for artificial intelligence evaluation for digital breast tomosynthesis evaluation, Clarity consortium for breast artificial intelligence, and DeepLook for artificial intelligence validation. All other authors declare no competing interests.

Acknowledgments

BK and BJE are partly supported by National Institutes of Health (NIH, grant number 1R21MD019360). JWG is a fellow of the 2022 Harold Amos Medical Faculty Development Program at the Robert Wood Johnson Foundation and declares support from the Radiological Society of North America Health Disparities grant (grant number EIHD2204), Lacuna Fund (grant number 67), NIH (National Institute of Biomedical Imaging and Bioengineering) Medical Imaging and Data Resource Center grant (under contracts 75N92020C00008 and 75N92020C00021), and National Heart, Lung, and Blood Institute (grant number R01HL167811). During the preparation of this work, the authors used large language models (Claude Opus 3-0) to make grammar and readability improvements. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication. All scientific content, literature analyses, interpretations, and conclusions are the original work of the authors.

References

- Achiam J, Adler S, Agarwal S, et al. Gpt-4 technical report. *arXiv* 2023; published online March 15. <http://arxiv.org/abs/2303.08774> (preprint).
- Singhal K, Tu T, Gottweis J, et al. Toward expert-level medical question answering with large language models. *Nat Med* 2025; **31**: 943–50.
- Saab K, Tu T, Weng W-H, et al. Capabilities of Gemini models in medicine. *arXiv* 2024; published online April 29. <http://arxiv.org/abs/2404.18416> (preprint).
- Podell D, English Z, Lacey K, et al. SDXL: improving latent diffusion models for high-resolution image synthesis. *arXiv* 2023; published online July 4. <http://arxiv.org/abs/2307.01952> (preprint).
- Ramesh A, Dhariwal P, Nichol A, Chu C, Chen M. Hierarchical text-conditional image generation with CLIP latents. *arXiv* 2022; published online April 13. <http://arxiv.org/abs/2204.06125> (preprint).
- Liu Y, Zhang K, Li Y, et al. Sora: a review on background, technology, limitations, and opportunities of large vision models. *arXiv* 2024; published online Feb 27. <http://arxiv.org/abs/2402.17177> (preprint).
- Chambon P, Bluethgen C, Delbrouck J-B, et al. RoentGen: vision-language foundation model for chest X-ray generation. *arXiv* 2022; published online Nov 23. <http://arxiv.org/abs/2211.12737> (preprint).
- Khosravi B, Li F, Dapamede T, et al. Synthetically enhanced: unveiling synthetic data's potential in medical imaging research. *EBioMedicine* 2024; **104**: 105174.
- Pinaya WHL, Tudosiu P-D, Dafflon J, et al. Brain imaging generation with latent diffusion models. *arXiv* 2022; published online Sept 15. <http://arxiv.org/abs/2209.07162> (preprint).
- Koetzier LR, Wu J, Mastrodicasa D, et al. Generating synthetic data for medical imaging. *Radiology* 2024; **312**: e232471.
- Giuffrè M, Shung DL. Harnessing the power of synthetic data in healthcare: innovation, application, and privacy. *NPJ Digit Med* 2023; **6**: 186.
- Jordon J, Szpruch L, Houssiau F, et al. Synthetic data—what, why and how? *arXiv* 2022; published online May 6. <http://arxiv.org/abs/2205.03257> (preprint).
- Yang R, Mandt S. Lossy image compression with conditional diffusion models. *Advances In Neural Information Processing Systems*. 2023. https://proceedings.neurips.cc/paper_files/paper/2023/file/ccf6d8b4a1fe9d9c8192f00c713872ea-Paper-Conference.pdf (accessed March 1, 2025).
- Unberath M, Zaech J-N, Gao C, et al. Enabling machine learning in X-ray-based procedures via realistic simulation of image formation. *Int J Comput Assist Radiol Surg* 2019; **14**: 1517–28.
- Guo X, Gichoya JW, Purkayastha S, Banerjee I. CVAD: an anomaly detector for medical images based on cascade VAE. In: Zamzmi G, Antani S, Bagci U, Linguraru MG, Rajaraman S, Xue Z, eds. *Medical image learning with limited and noisy data*. Lecture Notes in Computer Science, vol 13559. Springer, 2022: 187–96.
- Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks. *Commun ACM* 2020; **63**: 139–44.
- Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. In: *NIPS'20: Proceedings of the 34th International Conference on Neural Information Processing Systems*. Advances in Neural Information Processing Systems 2020: 6840–51.
- Xiao Z, Kreis K, Vahdat A. Tackling the generative learning trilemma with denoising diffusion GANs. *arXiv* 2021; published online Dec 15. <https://arxiv.org/abs/2112.07804> (preprint).
- Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* 2018; **321**: 321–31.
- Ktena I, Wiles O, Albuquerque I, et al. Generative models improve fairness of medical classifiers under distribution shifts. *Nat Med* 2024; **30**: 1166–73.
- Conte GM, Weston AD, Vogelsang DC, et al. Generative adversarial networks to synthesize missing T1 and FLAIR MRI sequences for use in a multisequence brain tumor segmentation model. *Radiology* 2021; **299**: 313–23.
- Rouzkroh P, Khosravi B, Faghani S, Moassefi M, Vahdati S, Erickson BJ. Multitask brain tumor inpainting with diffusion models: a methodological report. *arXiv* 2022; published online Oct 21. <http://arxiv.org/abs/2210.12113> (preprint).
- Khosravi B, Rouzkroh P, Erickson BJ, et al. Analyzing racial differences in imaging joint replacement registries using generative artificial intelligence: advancing orthopaedic data equity. *Arthroplast Today* 2024; **29**: 101503.
- Pérez-García F, Bond-Taylor S, Sanchez PP, et al. RadEdit: stress-testing biomedical vision models via diffusion image editing. *arXiv* 2023; published online Dec 20. <http://arxiv.org/abs/2312.12865> (preprint).

- 25 Khosravi B, Rouzrokh P, Mickley JP, et al. Few-shot biomedical image segmentation using diffusion models: beyond image generation. *Comput Methods Programs Biomed* 2023; **242**: 107832.
- 26 Rouzrokh P, Khosravi B, Mickley JP, Erickson BJ, Taunton MJ, Wyles CC. THA-net: a deep learning solution for next-generation templating and patient-specific surgical execution. *J Arthroplasty* 2024; **39**: 727–33.e4.
- 27 Yuan C, Duan J, Tustison NJ, Xu K, Hubbard RA, Linn KA. ReMiND: recovery of missing neuroimaging using diffusion models with application to Alzheimer's disease. *Imaging Neurosci* 2024; **2**: 1–14.
- 28 Kyung D, Kim J, Kim T, Choi E. Towards predicting temporal changes in a patient's chest X-ray images based on electronic health records. *arXiv* 2024; published online Sept 11. <http://arxiv.org/abs/2409.07012> (preprint).
- 29 Liu Q, Fuster-Garcia E, Hovden IT, et al. Treatment-aware diffusion probabilistic model for longitudinal MRI generation and diffuse glioma growth prediction. *IEEE Trans Med Imaging* 2025; published online Jan 23. <https://doi.org/10.1109/TMI.2025.3533038>.
- 30 Coyner AS, Chen JS, Chang K, et al. Synthetic medical images for robust, privacy-preserving training of artificial intelligence: application to retinopathy of prematurity diagnosis. *Ophthalmol Sci* 2022; **2**: 100126.
- 31 Shumailov I, Shumaylov Z, Zhao Y, Gal Y, Papernot N, Anderson R. The curse of recursion: training on generated data makes models forget. *arXiv* 2023; published online May 27. <http://arxiv.org/abs/2305.17493> (preprint).
- 32 Yang Q, Yan P, Zhang Y, et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Trans Med Imaging* 2018; **37**: 1348–57.
- 33 Kuanar S, Athitsos V, Mahapatra D, Rao KR, Akhtar Z, Dasgupta D. Low dose abdominal CT image reconstruction: an unsupervised learning based approach. In: 2019 IEEE International Conference on Image Processing (ICIP). Institute of Electrical and Electronics Engineers, 2019: 1351–55.
- 34 Heckel R, Jacob M, Chaudhari A, Perlman O, Shimron E. Deep learning for accelerated and robust MRI reconstruction. *MAGMA* 2024; **37**: 335–68.
- 35 Rueckel J, Huemmer C, Fieselmann A, et al. Pneumothorax detection in chest radiographs: optimizing artificial intelligence system for accuracy and confounding bias reduction using in-image annotations in algorithm training. *Eur Radiol* 2021; **31**: 7888–900.
- 36 Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 2004; **13**: 600–12.
- 37 Dolan B. Computer aided diagnosis in mammography: its development and early challenges. In: Fortieth Asilomar Conference on Signals, Systems and Computers. Institute of Electrical and Electronics Engineers, 2006: 821–25.
- 38 Ravuri S, Vinyals O. Classification accuracy score for conditional generative models. *Advances In Neural Information Processing Systems*. 2019. https://proceedings.neurips.cc/paper_files/paper/2019/hash/fcf55a303b71b84d326fb1d06e332a26-Abstract.html (accessed March 1, 2025).
- 39 Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A, Chen X. Improved techniques for training GANs. *arXiv* 2016; published online June 10. <http://arxiv.org/abs/1606.03498> (preprint).
- 40 Heusel M, Ramsauer H, Unterthiner T, Nessler B, Hochreiter S. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. *Advances in Neural Information Processing Systems*, 2017: 6629–40.
- 41 Bińkowski M, Sutherland DJ, Arbel M, Gretton A. Demystifying MMD GANs. *arXiv* 2018; published online Jan 4. <http://arxiv.org/abs/1801.01401> (preprint).
- 42 Zhou S, Gordon ML, Krishna R, Narcomey A, Fei-Fei L, Bernstein MS. HYPE: a benchmark for Human eYe Perceptual Evaluation of generative models. *Advances In Neural Information Processing Systems*. 2019. <https://proceedings.neurips.cc/paper/2019/file/65699726a3c601b9f31bf04019c8593c-Paper.pdf> (accessed March 1, 2025).
- 43 Liu Z, Wolfe S, Yu Z, et al. Observer-study-based approaches to quantitatively evaluate the realism of synthetic medical images. *Phys Med Biol* 2023; **68**: 074001.
- 44 Hessel J, Holtzman A, Forbes M, Le Bras R, Choi Y. CLIPScore: a reference-free evaluation metric for image captioning. *arXiv* 2021; published online April 18. <http://arxiv.org/abs/2104.08718> (preprint).
- 45 Li J, Li D, Xiong C, Hoi S. BLIP: bootstrapping language-image pre-training for unified vision-language understanding and generation. *Proceedings of the 39th International Conference on Machine Learning*. 2022. <https://proceedings.mlr.press/v162/li22n/li22n.pdf> (accessed March 1, 2025).
- 46 Dinh TM, Nguyen R, Hua B-S. TISE: bag of metrics for text-to-image synthesis evaluation. In: Avidan S, Brostow G, Cissé M, Farinella GM, Hassner T, eds. *Computer vision. Lecture Notes in Computer Science*. vol 13696. Springer, 2022: 594–609.
- 47 Priya RA, Patil AV, Bhende M, Thakare AD, Wagh S, eds. Object detection by stereo vision images. Wiley-Scrivener, 2022.
- 48 Osuala R, Skorupko G, Lazrak N, et al. *medigan*: a Python library of pretrained generative models for medical image synthesis. *J Med Imaging (Bellingham)* 2023; **10**: 061403.
- 49 Woodland M, Castelo A, Al Taie M, et al. Feature extraction for generative medical imaging evaluation: new evidence against an evolving trend. In: Linguraru MG, Dou Q, Feragen A, et al, eds. *Medical image computing and computer assisted intervention*. *Lecture Notes in Computer Science*, vol 15012. Springer, 2024: 87–97.
- 50 Jiang Y, Chen H, Loew M, Ko H. COVID-19 CT image synthesis with a conditional generative adversarial network. *IEEE J Biomed Health Inform* 2021; **25**: 441–52.
- 51 Zhang S, Xu Y, Usuyama N, et al. BiomedCLIP: a multimodal biomedical foundation model pretrained from fifteen million scientific image-text pairs. *arXiv* 2023; published online March 2. <http://arxiv.org/abs/2303.00915> (preprint).
- 52 Wu Q, Liu Y, Zhao H, et al. Uncovering the disentanglement capability in text-to-image diffusion models. *arXiv* 2022; published online Dec 16. <http://arxiv.org/abs/2212.08698> (preprint).
- 53 Yang T, Lan C, Lu Y, Zheng N. Diffusion model with cross attention as an inductive bias for disentanglement. *arXiv* 2024; published online Feb 15. <http://arxiv.org/abs/2402.09712> (preprint).
- 54 Kazerouni A, Aghdam EK, Heidari M, et al. Diffusion models in medical imaging: a comprehensive survey. *Med Image Anal* 2023; **88**: 102846.
- 55 Khosravi B, Rouzrokh P, Mickley JP, et al. Creating high fidelity synthetic pelvis radiographs using generative adversarial networks: unlocking the potential of deep learning models without patient privacy concerns. *J Arthroplasty* 2023; **38**: 2037–43.e1.
- 56 Lugmayr A, Danelljan M, Romero A, Yu F, Timofte R, Van Gool L. RePaint: inpainting using denoising diffusion probabilistic models. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Institute of Electrical and Electronics Engineers, 2022: 11461–71.
- 57 Linmans J, Raya G, van der Laak J, Litjens G. Diffusion models for out-of-distribution detection in digital pathology. *Med Image Anal* 2024; **93**: 103088.
- 58 Wolleb J, Bieder F, Sandkühler R, Cattin PC. Diffusion models for medical anomaly detection. In: Wang L, Dou Q, Fletcher PT, Speidel S, Li S, eds. *Medical image computing and computer assisted intervention*. *Lecture Notes in Computer Science*, vol 13438. Springer, 2022: 35–45.
- 59 Han T, Kather JN, Pedersoli F, et al. Image prediction of disease progression by style-based manifold extrapolation. *arXiv* 2021; published online Nov 22. <http://arxiv.org/abs/2111.11439> (preprint).
- 60 Akbar MU, Wang W, Eklund A. Beware of diffusion models for synthesizing medical images—a comparison with GANs in terms of memorizing brain MRI and chest x-ray images. *Mach Learn Sci Technol* 2025; **6**: 015022.
- 61 Somepalli G, Singla V, Goldblum M, Geiping J, Goldstein T. Understanding and mitigating copying in diffusion models. *arXiv* 2023; published online May 31. <http://arxiv.org/abs/2305.20086> (preprint).
- 62 Schwarz CG, Kremers WK, Therneau TM, et al. Identification of anonymous MRI research participants with face-recognition software. *N Engl J Med* 2019; **381**: 1684–86.

- 63 Macpherson MS, Hutchinson CE, Horst C, Goh V, Montana G. Patient reidentification from chest radiographs: an interpretable deep metric learning approach and its applications. *Radiol Artif Intell* 2023; 5: e230019.
- 64 Dockhorn T, Cao T, Vahdat A, Kreis K. Differentially private diffusion models. *arXiv* 2022; published online Oct 18. <http://arxiv.org/abs/2210.09929> (preprint).
- 65 Khosravi B, Mickley JP, Rouzrokh P, et al. Anonymizing radiographs using an object detection deep learning algorithm. *Radiol Artif Intell* 2023; 5: e230085.
- 66 Dathathri S, See A, Ghaisas S, et al. Scalable watermarking for identifying large language model outputs. *Nature* 2024; 634: 818–23.
- 67 Duan J, Kong F, Wang S, Shi X, Xu K. Are diffusion models vulnerable to membership inference attacks? Proceedings of the International Conference on Machine Learning. 2023. <https://proceedings.mlr.press/v202/duan23b/duan23b.pdf> (accessed March 1, 2025).
- 68 Moassefi M, Singh Y, Conte GM, et al. Checklist for reproducibility of deep learning in medical imaging. *J Imaging Inform Med* 2024; 37: 1664–73.
- 69 Alderman JE, Palmer J, Laws E, et al. Tackling algorithmic bias and promoting transparency in health datasets: the STANDING Together consensus recommendations. *Lancet Digit Health* 2025; 7: e64–88.
- 70 Tang R, Liu L, Pandey A, et al. What the DAAM: interpreting stable diffusion using cross attention. *arXiv* 2022; published online Oct 10. <http://arxiv.org/abs/2210.04885> (preprint).
- 71 Horwitz E, Hoshen Y. Confusion: confidence intervals for diffusion models. *arXiv* 2022; published online Nov 17. <http://arxiv.org/abs/2211.09795> (preprint).
- 72 Luccioni AS, Akiki C, Mitchell M, Jernite Y. Stable bias: analyzing societal representations in diffusion models. *arXiv* 2023; published online March 20. <http://arxiv.org/abs/2303.11408> (preprint).
- 73 Seyyed-Kalantari L, Zhang H, McDermott MBA, Chen IY, Ghassemi M. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. *Nat Med* 2021; 27: 2176–82.
- 74 Drukker K, Chen W, Gichoya J, et al. Toward fairness in artificial intelligence for medical image analysis: identification and mitigation of potential biases in the roadmap from data collection to model deployment. *J Med Imaging (Bellingham)* 2023; 10: 061104.
- 75 Kitamura FC, Prevedello LM, Colak E, et al. Lessons learned in building expertly annotated multi-institution datasets and hosting the RSNA AI challenges. *Radiol Artif Intell* 2024; 6: e230227.
- 76 Zhu Z, Zhao H, He H, et al. Diffusion models for reinforcement learning: a survey. *arXiv* 2023; published online Nov 2. <http://arxiv.org/abs/2311.01223> (preprint).
- 77 Subtle Medical. Subtle Medical's SubtleHD™ wins FDA clearance, setting a new benchmark for MRI image quality and speed. Feb 14, 2025. <https://subtlemedical.com/subtle-medicals-subtlehd-wins-fda-clearance-setting-a-new-benchmark-for-mri-image-quality-and-speed/> (accessed April 3, 2025).

© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

STATE-OF-THE-ART REVIEW

Atherosclerotic Cardiovascular Disease Risk Prediction Models in China, Japan, and Korea



Implications for East Asians?

Patricia K. Nguyen, MD, MBA,^a Dong Zhao, MD, PhD,^b Tomonori Okamura, MD, PhD,^c Hyeon Chang Kim, MD, PhD,^d Nathan D. Wong, PhD,^e Eugene Yang, MD, MS^f

ABSTRACT

The management of atherosclerotic cardiovascular disease (ASCVD) in the United States is currently based upon large epidemiological studies in primarily non-Hispanic White subjects. Although this strategy provides a uniform approach that is simpler to implement, it may result in inappropriately targeting certain Asian populations for treatment based on inaccurate ASCVD risk estimation. In this state-of-the-art review, we detail the similarities and differences in the prevalence of ASCVD and its risk factors among Chinese, Japanese, and Korean people living in the United States and in their native countries. We highlight the limitations of current risk calculators when applied to East Asian immigrants and summarize risk stratification approaches in China, Japan, and Korea. Our review underscores the need to disaggregate registry, cohort, and clinical trial data by East Asian subgroups, to actively engage these populations in research, and to initiate studies to better define ASCVD risk in East Asian people living in the United States. (JACC Asia. 2025;5:333-349) © 2025 The Authors. Published by Elsevier on behalf of the American College of Cardiology Foundation. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

East Asian people make up 20.7% of the world's population based on the latest estimates by the United Nations.¹ As the fastest growing immigrant population in the United States, Asians comprise approximately 7% of the U.S. population, with East Asians making up the largest Asian subgroup (eg, ~40%). In 2019, cardiovascular disease (CVD) claimed 5.2 million lives in East Asian countries.² Atherosclerotic cardiovascular disease (ASCVD), including ischemic heart disease, ischemic stroke, and peripheral arterial disease, was the leading cause of cardiovascular morbidity and mortality

among East Asian persons.² Declining fertility rates and increasing life expectancies will ensure that ASCVD continues to be one of the most common non-communicable, chronic diseases affecting East Asian persons for future decades.³

Notably, epidemiological characteristics of ASCVD vary among East Asian persons compared with Southern, Western, and Central Asian persons.² These differences exist even within East Asian subpopulations. The proportional mortality rate of CVD is as low as 25% in the Japanese and South Korean populations but as high as 40% in Chinese people,

From the ^aDepartment of Medicine, Division of Cardiovascular Medicine, Stanford University, Stanford, California, USA; ^bDepartment of Epidemiology, Beijing Anzhen Hospital, Capital Medical University, Beijing Institute of Heart, Lung, and Blood Vessel Diseases, National Clinical Research Center for Cardiovascular Diseases, Beijing, China; ^cDepartment of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan; ^dDepartment of Preventive Medicine, Integrative Research Center for Cerebrovascular and Cardiovascular Diseases, Yonsei University College of Medicine, Seoul, Korea; ^eHeart Disease Prevention Program, Division of Cardiology, University of California, Irvine, Irvine, California, USA; and the ^fDivision of Cardiology, Department of Medicine, University of Washington School of Medicine, Seattle, Washington, USA. The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the [Author Center](#).

Manuscript received September 24, 2024; revised manuscript received January 2, 2025, accepted January 13, 2025.

ABBREVIATIONS AND ACRONYMS

ACC = American College of Cardiology

ASCVD = atherosclerotic cardiovascular disease

AHA = American Heart Association

BMI = body mass index

CAC = coronary artery calcium

CHD = coronary heart disease

CVD = cardiovascular disease

HDL-C = high-density lipoprotein cholesterol

KSH = Korean Society of Hypertension

LDL-C = low-density lipoprotein cholesterol

PCE = Pooled Cohort Equation

PR = prevalence ratio

TC = total cholesterol

highlighting the need for both targeted and personalized, therapeutic strategies for East Asian subgroups.²

Whether the observed differences in the epidemiological characteristics in East Asian subgroups living in Asia extend to East Asian Americans is unknown. Because the number of East Asian immigrants living in the United States is growing rapidly, this question demands further investigation. In 2019 alone, 4.2 million East Asians immigrated to the United States, a sharp rise from the 250,000 who immigrated in 1960.⁴ Similar to those living in Asia, ASCVD is a major cause of morbidity and mortality. ASCVD risk varies widely between Asian American subgroups,⁵⁻⁷ underscoring the need to disaggregate data to refine risk assessment and optimize CVD management in East Asian Americans.

Cardiovascular clinical trials inclusive of East Asian Americans are limited. Like other Asian subgroups,⁸ East Asian U.S. immigrants and their offspring are not adequately represented among major U.S. prospective CVD cohort studies, which form the basis of the American College of Cardiology (ACC)/American Heart Association (AHA) ASCVD Pooled Cohort Equation (PCE) as well as other risk calculators. Although the recently developed AHA PREVENT (American Heart Association Predicting Risk of cardiovascular disease EVENTS) risk calculator, derived from a meta-analysis of 25 data sets, has removed race/ethnicity altogether arguing their effects may be already reflected in socioeconomic data,⁹ the accuracy of this calculator has not been specifically validated in East Asian persons. As a result, current risk stratification models are not adequate to predict the development of ASCVD in East Asian Americans. Whether ASCVD risk assessment tools based on long-term cohort studies conducted in East Asian countries can accurately predict risk in East Asian Americans is unknown. Although recalibration can create a useful ASCVD prediction model and has been proposed to improve accuracy, it is preferable to develop and validate a model using local data if available. Even among countries classified with similar risk levels, there are considerable differences in incidence rates of CVD. For example, in Korea, SCORE2-AP recalibrated SCORE2 to the Asia-Pacific region and categorized risk into 4 groups (low, moderate, high, and very-high risk). Korea is included in the low-risk group, but this low-risk regional model also overestimates risk in Korea.¹⁰ Additionally, the distribution of CVD subtypes varies by region and period, so the coefficients may differ as

well. For example, although the coefficients for each predictor are somewhat similar, regional differences between models still exist. Thus, risk calculators that include the most recent local data would be the most appropriate. These knowledge gaps highlight the need for a more thorough understanding of variations in ASCVD disease burden and cardiovascular risk factor prevalence among East Asian people residing in the United States to improve risk stratification and management strategies.

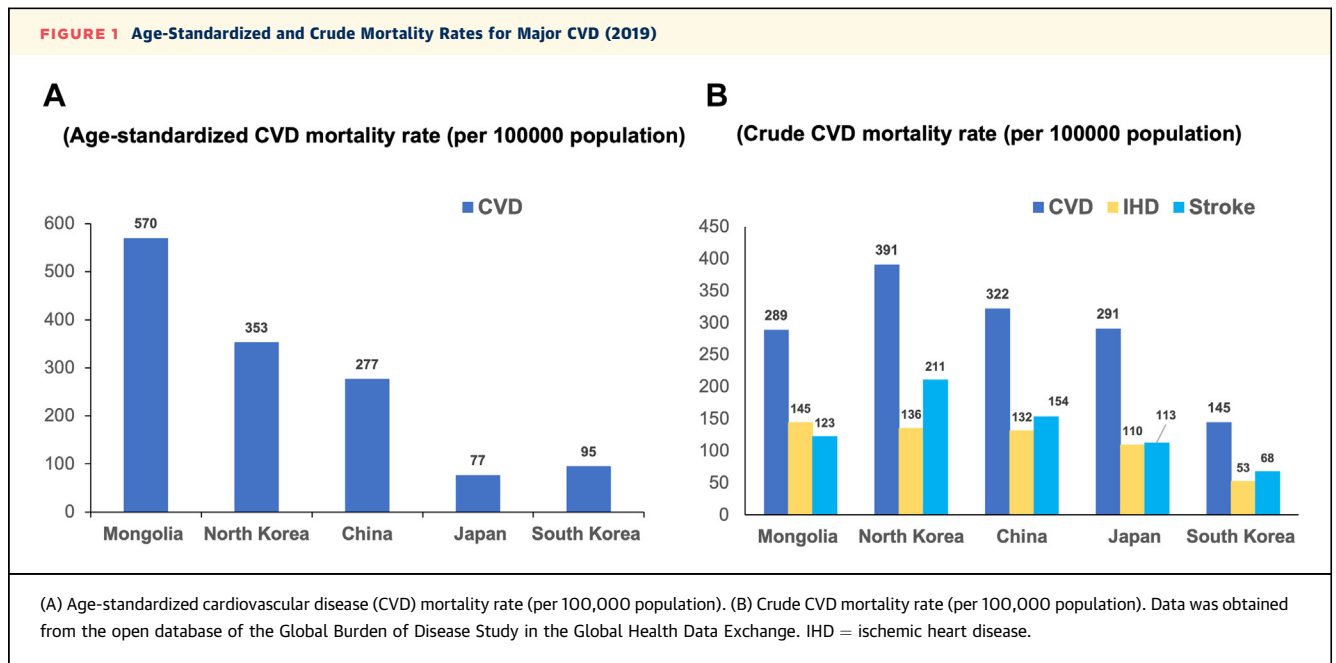
In this review, we highlight the similarities and differences in the epidemiology, diagnosis, and treatment of ASCVD for individuals of East Asian origin who immigrated to the United States and their offspring (“East Asian Americans”) compared with those living in East Asia (“East Asian natives”). We identify major knowledge gaps in our understanding of ASCVD risk and explore opportunities and strategies to close these gaps through future clinical and research initiatives. Because of the lack of data and smaller size of other East Asian countries, this review focuses on studies performed in East Asian persons living or originating from China, Japan, and the Republic of Korea (South Korea).

ASCVD PREVALENCE AND RISK FACTORS IN EAST ASIAN POPULATIONS

DEFINITION OF EAST ASIAN POPULATIONS.

Based on definitions provided by the United Nations, the geographic boundaries of the Asia-Pacific region include Turkey in the west, the Pacific Island of Kiribati in the east, the Russian Federation in the north, to New Zealand in the south.¹¹ Located in the Asia-Pacific region, major East Asian countries include the People’s Republic of China (China, including the special administrative regions of Hong Kong, Macao, and Taiwan), Democratic People’s Republic of Korea (North Korea), Republic of Korea (South Korea), Mongolia, and Japan. The major ethnic groups in East Asia include Han, Korean, and Yamato. Minor groups include Bai, Hui, Tibetan, Turkic, Manchus, Ryukyuan, Ainu, Zhuang, Mongols, and many other groups. In 2023, approximately 1.66 billion people live in East Asia, comprising 22% of the world’s population.¹

Additionally, millions of East Asian immigrants and their offspring reside outside of the geographical boundaries of East Asia. Since the mid-1960’s, the number of immigrants from East Asia to the United States has risen sharply. According to calculations by the Migration Policy Institute of U.S. Census Data,⁴ between 2000 and 2019, the number of East Asian persons living in the United States has increased by



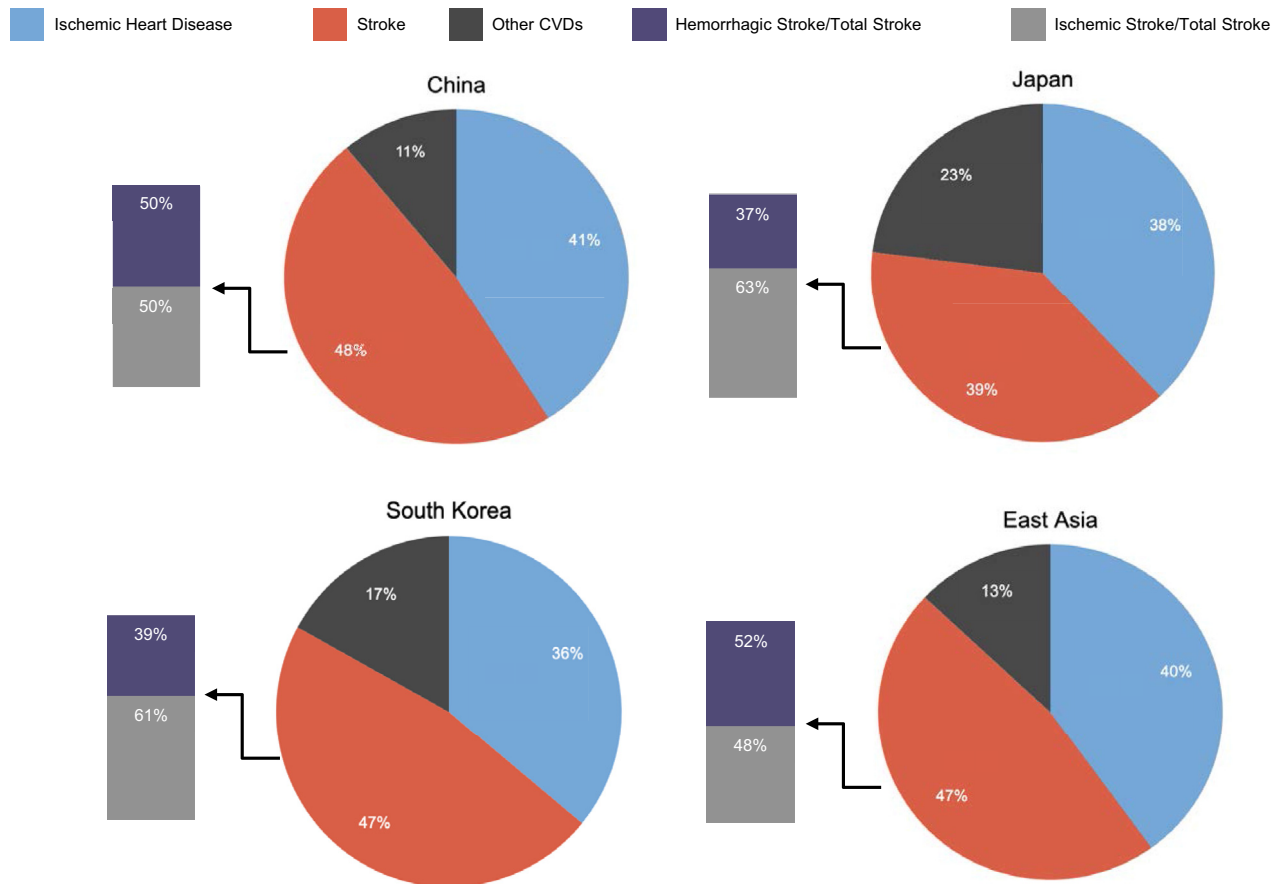
81% in the last 2 decades, outpacing Hispanic and non-Hispanic Black persons, whose numbers grew by 70% and 20%, respectively, during the same time frame. Immigrants from China and Korea accounted for 19% and 7% of the total Asian immigrant population, respectively.⁴

EPIDEMIOLOGY OF ASCVD IN EAST ASIAN POPULATIONS LIVING IN ASIA AND IN THE UNITED STATES. An aging population together with the exponential rise in obesity and diabetes will continue to fuel the higher incidence of ASCVD among East Asian Americans and East Asian natives. Below, we summarize and compare the existing data on the epidemiology of ASCVD in East Asian natives and East Asian Americans.

It is well established that East Asian countries have a specific epidemiological pattern of CVD. In 2019, nearly 5.2 million East Asian natives died of CVD. Overall crude CVD mortality in East Asia was 349 of 100,000 with notable differences in mortality rates across East Asian countries (Figure 1).² Of the 5 major countries of this region, South Korea had the lowest crude CVD mortality rate (145 of 100,000) while North Korea had the highest (391 of 100,000). In 2019, ischemic heart disease and ischemic/hemorrhagic stroke accounted for approximately 87% of all CVD deaths in East Asia with stroke making up more than one-half of them (Figure 2).² Of total deaths in East Asian countries, Japan had the lowest proportion of stroke deaths (39%), while China (48%) and South

Korea (47%) had comparable rates. Although hemorrhagic strokes appeared to be slightly more frequent in East Asia overall, the proportion of hemorrhagic strokes varied significantly across the region. On average, 52% of stroke deaths were attributed to hemorrhage stroke, ranging from as low as 36% in Japan to as high as 50% in China.² Although peripheral vascular disease continues to be an important contributor to CVD, the availability of peripheral vascular disease mortality data in East Asian countries remains limited.

National-level epidemiological data on ASCVD in East Asian Americans is limited, forcing us to rely on aggregated data and single center studies to estimate ASCVD prevalence and mortality rates among East Asian Americans. Based on the 2018 NHIS (National Health Interview Survey) conducted by the Centers for Disease Control and Prevention, the age-adjusted prevalence of stroke among persons 18 years of age or older was similar between non-Hispanic Asian (2.7%) and non-Hispanic White (2.7%) persons, whereas the prevalence of coronary heart disease (CHD) was lower in Non-Hispanic Asian persons at 4.4% compared with 5.7% in Non-Hispanic White persons.¹² Based on the 2019 National Vital Statistics Report conducted by the Centers for Disease Control and Prevention, the age-adjusted death rate for cerebrovascular disease was lower in non-Hispanic Asian vs non-Hispanic White persons (cerebrovascular disease: 29.3 of 100,000 vs 35.7 of 100,000).¹³ Greater differences were observed in heart disease

FIGURE 2 Proportion of Subtypes of CVD in Total CVD Death

Percentage of total cardiovascular disease (CVD) deaths attributable to ischemic heart disease, stroke, and other CVDs in China, Japan, and South Korea, and all of East Asia. The proportion of deaths caused by stroke are further stratified into the percentage of total stroke deaths caused by ischemic or hemorrhagic stroke. Data were obtained from the open database of the Global Burden of Disease Study in the Global Health Data Exchange.

mortality rates for non-Hispanic Asian persons compared with non-Hispanic White persons. The age-adjusted heart disease death rate was more than 2-fold lower in non-Hispanic Asian persons (mean 79.2 per 100,000 with 101.6 per 100,000 in men and 61.6 per 100,000 in women) vs non-Hispanic White persons (mean 166.4 per 100,000; 210.7 per 100,000 in men, and 129.6 per 100,000 in women).

The data highlighted in the previous text includes data for all Asian persons living in the United States (including those of South Asian origin that often exhibit very different risk profiles from East Asians), and the exact prevalence of ASCVD and its associated mortality in East Asian Americans is uncertain. Most surveys lack or have limited data disaggregating East Asian subgroups from all Asian and Pacific Islander

persons. Although the U.S. Census Bureau has made great strides to ensure accurate representation of Asian immigrants in surveys through outreach programs, native-language interviews, subgroup categorization, and oversampling, limited disaggregated data is currently available for East Asian Americans from national public surveys such as the NHIS and NHANES (National Health and Nutrition Examination Survey). Generalizability of these data are challenging because some surveys require a minimum level of English proficiency.

Several investigators have published disaggregated data in East Asian immigrant subgroups. In an early cross-sectional study of 21,722 Asians living in Northern California from 2007 to 2009 and receiving care from a private health care system, the prevalence

of CHD and stroke among East Asian persons (ie, Chinese, Japanese, and Korean) was 2.07% (274 of 13,181) and 0.75% (100 of 13,181), respectively. These rates were substantially lower than the prevalence of CHD (3.9%, 2,805 of 72,701) and stroke (1.2%, 846 of 72,701) among non-Hispanic White persons.¹⁴ A more recent 2020 study that included 51,006 East Asian persons from Kaiser Permanente Northern California with CVD risk factors, the incidence of CHD (including coronary revascularization, MI, and cardiovascular death) was 4.46% in East Asian persons compared with 6.29% in non-Hispanic White persons over a 10-year follow-up period.¹⁵

Three studies have examined mortality rates of ischemic heart disease and stroke for the 6 largest Asian American subgroups, including East Asian immigrants from China, Japan, and Korea and 3 other Asian subgroups (Asian Indian, Filipino, and Vietnamese).¹⁶⁻¹⁸ Two of the earlier studies used mortality data from 34 states that reported deaths for each Asian subgroup in the 2003 to 2010 Multiple Cause of Death Mortality Database from the National Center for Health Statistics.^{16,17} A more recent study by Shah et al¹⁸ incorporated 2017 death rates and found that age-adjusted mortality rates from ischemic heart disease are lower for non-Hispanic Asian persons compared with Hispanic and non-Hispanic White persons; however, age-adjusted mortality rates from cerebrovascular disease are comparable between Asian and non-Asian persons. Analysis of Asian subgroups reveals significant differences in mortality rates with Korean persons having the highest mortality rates among the East Asian population but much lower than South Asian persons. More contemporary data on the incidence, prevalence, and mortality associated with ASCVD for Chinese, Japanese, and Korean Americans as well as other East Asian subgroups are critical to refinement of risk assessment tools and treatment strategies.

PREVALENCE OF ASCVD RISK FACTORS IN EAST ASIAN POPULATIONS LIVING IN ASIA AND IN THE UNITED STATES. Cardiovascular risk factors, including hypertension, diabetes, dyslipidemia, tobacco use, overweight/obesity, and lifestyle factors such as unhealthy nutritional practices and physical inactivity, contribute to the development of ASCVD. Understanding the prevalence of major CVD risk factors among East Asian immigrants is critical for predicting the risk of developing ASCVD. Accurate and current information on the prevalence of cardiovascular risk factors in East Asian immigrants is not only the basis for evaluation of which risk prediction models are most accurate for East Asians, but also

form the basis for recalibration of these models. The following sections summarize existing data on the prevalence of cardiovascular risk factors for East Asian people living in their native countries and in the United States, respectively.

The epidemic of unhealthy lifestyles continues to drive the prevalence of ASCVD risk factors in East Asia, as evidenced by the data collected by the NCD Risk Factor Collaboration (NCD-RisC),¹⁶ a network of health scientists that provides rigorous and timely global data on major risk factors for non-communicable diseases. According to the most recent data available, South Korea had the lowest age-standardized prevalence of hypertension (13.8% in men and 8% in women) and obesity (4.5% in men and 5.0% in women). Japanese people had the highest mean levels of total cholesterol (TC) (4.98 mmol/L in both men and women) and high-density lipoprotein-cholesterol (HDL-C) (1.46 mmol/L in men and 1.77 mmol/L in women), while Chinese people had the lowest mean levels of TC (4.56 mmol/L in men and 4.66 mmol/L women) and HDL-C (1.20 mmol/L in men and 1.35 mmol/L in women). The U.S. population had much higher prevalence of obesity (36.0% in men and 38.1% in women) compared with the risk factor profiles of East Asian countries.^{14,19} Unfortunately, the numbers reported in the NCD-RisC reflect percentages in the United States as an entire population, which limit our ability to compare East Asian persons and their U.S. counterparts and further highlight the need for a global effort to disaggregate data by race and ethnicity. Moreover, the aforementioned percentages of obesity defined in the United States reflect body mass index (BMI) cutoffs for obesity defined in non-Asian persons. Asian American people are more likely to have central obesity, which remains a major risk factor for the development of type 2 diabetes, metabolic syndrome, and CVD at lower BMIs than non-Asian people. Evidence supporting racial differences in fat distribution between Asian and non-Asian people has prompted the lowering of the cutoff for obesity for Asian persons to ≥ 27.5 and ≥ 25.0 kg/m² rather than the cutoff 30 kg/m² for non-Asian persons by the World Health Organization and other regulatory bodies (eg, Japanese Society for the Study of Obesity, Korean Society for the Study of Obesity),²⁰ respectively. Thus, the estimates of obesity by the NCD-RisC may underestimate the true incidence of obesity in East Asian Americans and non-East Asian Americans.

The most recent estimates of cardiovascular risk factor prevalence in Asian Americans were published by Commodore-Mensah et al²⁰ and Koirala et al,⁸ who independently evaluated the heterogeneity of CVD

risk factors among U.S. Asian immigrants using a cross-sectional analysis collected from NHIS between 2010 to 2018. They compared the prevalence of CV risk factors between 474,968 White adults and 33,973 Asian immigrants including an estimated 9,753 individuals from the Asia region (including Asia Minor, China, Democratic People's Republic of Korea [North Korea], Japan, Mongolia, and Republic of Korea [South Korea]). The remaining Asian immigrants included in the analysis were from South Asia (Afghanistan, Bangladesh, Bhutan, India, Nepal, Pakistan, Sri Lanka) and Southeast Asia (Burma, Cambodia, Indonesia, Laos, Malaysia, Philippines, Singapore, Thailand, and Vietnam) regions. Notably, the geoscheme for Asia used by the authors differs from the geoscheme defined by the United Nations and adopted by the writers of this review.

Compared with non-Hispanic White persons, East Asian immigrants were less likely to have hypertension (20.8% vs 34.3%), diabetes (6.3% vs 9.6%), dyslipidemia (21.9% vs 31.7%), or overweight/obesity (47.1% vs 64.1%). Rates of smoking were not significantly different between non-Hispanic white and East Asian Americans. Compared with South Asian immigrants, East Asian immigrants were more likely to have hypertension (20.8% vs 16.6%), but less likely to report having diabetes (6.3% vs 8.6%) or overweight/obese (47.1% vs 71.54%). The rate of dyslipidemia between East Asian immigrants and South Asian immigrants was comparable (21.9% vs 20.8%). Compared with Southeast Asian immigrants, East Asian immigrants were less likely to have hypertension (20.8% vs 31.2%), diabetes (6.30% vs 11.3%), dyslipidemia (21.9% vs 33.41%), or overweight/obesity (47.1% vs 59.8%). Physical inactivity was equally present in all Asian subgroups with over 50% of patients not meeting physical activity guidelines. Compared with non-Hispanic White persons, East Asian persons were more likely to be physically inactive, defined by the proportion of the population not meeting ACC/AHA guidelines for exercise and physical activity (50.8% vs 48.3%). Even after adjustment for age, sex, and socioeconomic factors, East Asian persons were more likely to report physical inactivity (prevalence ratio [PR] 1.14, 95% CI: 1.09-1.19) but less likely to report hypertension (PR 0.72, 95% CI: 0.67-0.77), overweight/obesity (PR 0.83, 95% CI: 0.80-0.87), diabetes mellitus (PR 0.80, 95% CI: 0.70-0.91), high cholesterol (PR 0.83; 95% CI: 0.7700-0.90), and current smoking (PR 0.53, 95% CI: 0.46-0.60) than their non-Hispanic White counterparts. Unfortunately, this analysis was unable disaggregate subjects by their country of origin—a critical variable to estimate the incidence







and prevalence of CVD risk factors among East Asian American subgroups and essential for creating a foundation for accurate risk estimation.

THE IMPACT OF ACCULTURATION AND ENVIRONMENTAL EFFECTS OF ASCVD RISK PROFILES. It is fundamental to understand the potential differences in CVD risk factors and ASCVD mortality profiles of East Asian natives and East Asian Americans before we can develop an accurate risk assessment strategy. To better account for acculturation and environmental effects, risk factor tools for East Asian immigrants may need to consider the impact of immigration history and generational status on risk factor profiles. Few studies have directly compared the profiles of CVD risk factors between Asian Americans and Asian immigrants. In the REACH (REduction of Atherothrombosis for Continued Health) registry that enrolled close to 3,000 Chinese patients with atherothrombotic disease living outside and inside mainland China, rates of obesity, hypertension, hypercholesterolemia, and diabetes were lower in mainland Chinese people than those in China Hong Kong, China Taiwan, Singapore, Western Europe, and North America.²¹ In a recent cross-sectional analysis of the Japanese population, Hirooka et al²² showed that Japanese Americans reported less sedentary behavior but were also less likely to engage in physical activity and formal exercise. Although the findings are based on studies conducted a decade ago, they highlight the need to account for environmental and cultural factors to create an accurate risk stratification tool for East Asian Americans. A recent study found that acculturation was associated with a heterogeneous pattern of CVD risk factors among Asian American subgroups, highlighting the need for more research to better understand these differences and to guide culturally concordant interventions.²³

CURRENT STATE OF ASCVD RISK CALCULATORS FOR EAST ASIAN POPULATIONS

In the 2018 ACC/AHA Guideline on the Management of Blood Cholesterol,²⁴ the PCE was used for risk assessment for primary prevention to guide the eligibility for statin therapy. It was developed by combining 4 U.S. community cohort studies, including the ARIC (Atherosclerosis Risk in Communities), CHS (Cardiovascular Health Study), CARDIA (Coronary Artery Risk Development in Young Adults), and Framingham Study. It predicts the 10-year risk of ASCVD (including nonfatal and fatal CHD and stroke). The risk factors used in this model include sex, race (non-Hispanic White, African American, or other),

CENTRAL ILLUSTRATION East Asian Cardiovascular Risk Calculators

Risk Factors	China 	Japan 	Korea* 	ACC/AHA PCE 	AHA PREVENT 	ESC SCORE2 
Alcohol use	○	◐	◐	○	○	○
Blood pressure	●	●	●	●	●	●
Body mass index (BMI)	◐	◐	◐	○	●	○
Cholesterol	●	●	●	●	●	●
Chronic kidney disease	◐	◐	◐	●	●	●
Diabetes mellitus	●	●	●	●	●	●
Family history of CAD	◐	○	◐	●	○	●
Physical activity	○	◐	◐	●	○	●
Tobacco use	●	●	●	●	●	●

● = Included in all risk calculators; ◐ = Included in some risk calculators; ○ = Not included in risk calculators;
 ● = Risk Enhancer (ACC/AHA) or Risk Modifier (ESC)
 * = Korea denotes the Republic of South Korea

Nguyen PK, et al. *JACC Asia*. 2025;5(3):333-349.

Comparing cardiovascular risk calculators: China, Japan, Korea, United States, and Europe. Filled circles: included in all risk calculators; half-filled circles: included in some risk calculators; open circles: not included in risk calculators; circles with bold outlines: risk enhancer (American College of Cardiology [ACC]/American Heart Association [AHA]) or risk modifier (European Society of Cardiology [ESC]). *Korea denotes the Republic of South Korea. CAD = coronary artery disease.

age, systolic blood pressure, treatment for hypertension, TC, HDL-C, smoking, and diabetes. Notably, the cohorts used in the PCE included very few Asian subjects. Moreover, these cohorts were initiated in the 1980s or earlier when CVD risks were higher or have tended to overestimate ASCVD risks caused by this secular effect. Also, as detailed in the following text, estimated 10-year ASCVD risk is generally lower in Asian American persons compared with non-Hispanic White persons. Accordingly, Asian-specific risk stratification calculators linked with clinical guidelines are needed to bridge these gaps (**Central Illustration**).¹⁸

ASCVD RISK PREDICTION IN CHINA. Several risk scores have been evaluated for ASCVD risk prediction in China. In 2003, the CMCS (Chinese Multi-provincial Cohort Study) of 30,121 adults (ages 35 to 64 years) evaluated the performance of the Framingham CHD risk score with equations derived from CMCS.²⁵ The investigators found that when compared directly and after recalibration, the original Framingham equation

significantly overestimated absolute CHD risk in the CMCS cohort, which was mainly driven by differences in the mean CHD risk and the levels of major risk factors between the 2 cohorts. Specifically, the 10-year CHD event rates were 8.0% and 2.8% in Framingham men and women, respectively, compared with 1.5% and 0.6% in the CMCS men and women. With respect to cardiac risk factors, men and women in the Framingham cohort had higher TC and lower HDL-C, respectively, than those in the CMCS cohort. However, men in the Framingham cohort had lower and women had higher smoking rates in comparison with men and women in the CMCS cohort. Importantly, a recalibration of the Framingham Risk Score (FRS) equation (ie, the mean risk of CHD and mean levels of risk factors derived from Framingham cohort), performed by replacing the corresponding estimations from CMCS cohort, substantially improved the accuracy of risk prediction.²⁵ Even after recalibrating Framingham risk calculators, the China-MUCA (Multicenter Collaborative Study of Cardiovascular Epidemiology)

TABLE 1 ASCVD Risk Assessment Tools in China

First Author	Study Population	Predictors (Other Than Age and Sex)	Outcomes	Duration of Risk Prediction	Validation	Risk Tools
Liu et al ²⁵	30,121 men and women aged 35-64 y	SBP, TC, HDL-C, smoking status, diabetes	CHD	10 y	Internal	Equations
Wu et al ²⁶	11,336 men and women aged 35-59 y	SBP, TC, smoking status, diabetes, BMI	ASCVD	10 y	Internal External	Risk scores
Yang et al ²⁷ and Xing, et al ³¹	27,020 men and women aged 35-74 y	SBP, TC, HDL-C, smoking, diabetes, WC, geographic region, urban/rural, family history of ASCVD	ASCVD Stroke	10 y Lifetime	Internal External	Equations and web-based risk calculator
Wang et al, ²⁸ Wang et al, ³⁰ and Wang et al ²⁹	30,121 men and women aged 35-64 y	SBP, LDL-C, HDL-C, smoking status, diabetes	ASCVD	10 y Lifetime	Internal	Equations, risk charts, and flowcharts of risk classification

ASCVD = atherosclerotic cardiovascular disease; BMI = body mass index; CHD = coronary heart disease; HDL-C = high-density lipoprotein cholesterol; LDL-C = low-density lipoprotein cholesterol; SBP = systolic blood pressure; TC = total cholesterol; WC = waist circumference.

cohort study reported that CHD risk for both men and women was still substantially overestimated.²⁶ More recently, another large cohort study, China-PAR (Prediction for ASCVD Risk in China) project, found that the PCE had low discrimination ability and poor calibration for Chinese men.²⁷ These findings highlighted the importance of developing CVD risk prediction models based on data from China cohort studies.

Based on data from the CMCS cohort study, the first sex-specific ASCVD risk prediction equations and stratification algorithms were published in 2003 and subsequently updated in 2018.^{28,29} The China-MUCA study and China-Par project developed and published risk predictive models to estimate 10-year risk ASCVD in Chinese people.^{26,27} A comparison of these risk prediction models is shown in **Table 1**. After the development of 10-year ASCVD risk equations from large, long-term cohort studies, CMCS and China-PAR cohort studies were utilized to create lifetime ASCVD risk prediction models for young and middle-aged people.^{30,31} The lifetime ASCVD risk prediction tool can identify those with lower 10-year ASCVD risk but higher lifetime risk, thereby, facilitating earlier prevention intervention, including motivating lifestyle modifications for these individuals. All of these studies developed either categorized algorithms of risk classification by flowcharts, risk scoring systems, or web-based risk calculators to facilitate the application of risk assessment into clinical practice and for patient education.²⁶⁻²⁸

Since 1999, ASCVD risk stratification-based clinical decision making has been recommended in China by relevant CVD prevention practice guidelines to inform treatment strategies and targets for risk factor control.³²⁻³⁴ The 1999 Chinese guidelines for the management of hypertension was the first to employ

risk assessment to guide decision-making for blood pressure treatment.³² The risk classification criteria recommended by the World Health Organization/International Society of Hypertension blood pressure guideline were used because no risk assessment tools were available in China at that time. The 2007 Chinese guidelines on prevention and treatment of dyslipidemia in adults was the first to use ASCVD risk classification based on risk prediction models developed from Chinese cohort studies (CMCS and China-MUCA) and was notable for its heavy weighting of hypertension.³³ Risk classification charts were updated based on new risk prediction equations and lifetime risk estimation developed from CMCS data for the 2016 Chinese management of dyslipidemia guideline for adults and 2017 China guideline for CVD prevention.^{35,36}

In 2020, the Chinese Guideline on the Primary Prevention of Cardiovascular Diseases further modified their recommendations for risk classification by adding chronic kidney disease stage 3/4 to diabetes, age >40 years, and very high low-density lipoprotein-cholesterol (LDL-C) (≥ 4.9 mmol/L) as high-risk conditions.³⁴ Ten risk-enhancing factors were added, including coronary artery calcium (CAC) score ≥ 100 Agatston units. CAC is well-established as a risk stratifying tool among Chinese adults from a large study from China of 4,425 patients with suspected coronary disease scanned for presence of CAC and coronary plaque.³⁷ At 3 years of follow-up, the risk of major adverse cardiac events increased with higher CAC scores (CAC 0, 2.1%; CAC 1 to 100, 12.9%; CAC 101 to 400, 16.3%; and CAC >400, 33.8%; log-rank $P < 0.001$). The C-statistics improved with addition of CAC and plaque characteristics over traditional risk factors alone: 0.71 for clinical risk factors, which improved to 0.82 by adding CAC and further

improved to 0.93 by adding coronary computed tomography angiographic (CTA) plaque information (both $P < 0.001$). In addition, the U.S. Multiethnic Study of Atherosclerosis showed CAC strongly predicted future ASCVD events for Chinese American participants (as well as other racial/ethnic groups) over 10 years, with those having CAC scores >100 identified to have an ASCVD risk $\geq 7.5\%$ making them eligible for initiation of statin therapy.³⁸ Recently, a new risk classification system for secondary prevention was issued by the Chinese Society of Cardiology in an expert consensus statement. Instead of considering all patients with ASCVD as a very high-risk group in previous guidelines, this statement further identified a subset at extreme ASCVD risk for more intensive lipid lowering treatment.³⁹

Although country-specific risk assessment tools are readily available in China, significant challenges remain. Most of the cohort studies collected baseline information in the 1990s, but there have been changes in risk factor prevalence among more contemporary target populations. Risk prediction models need to be recalibrated, and some studies have started to evaluate these older equations with data collected from large contemporary cohorts. For instance, Liu et al⁴⁰ recently applied the PCE and China-PAR algorithms to an electronic health record cohort of 226,406 participants and showed that although both models had good discrimination, the China-PAR model substantially underpredicted risk, and the PCE overpredicted risk in men and underpredicted risk in women, although recalibration improved this. Additionally, although most of the risk enhancing factors increased ASCVD risk in observational studies, few studies in China showed that these risk-enhancing factors, except for CAC score, could significantly improve performance of risk calculators. Several new risk factors have been considered as potential contributors to residual risk in Chinese people, including small density LDL, cholesterol-overloaded HDL particles,⁴¹ and higher circulating levels of PCSK9.⁴⁰ However, the practical use of these new risk factors in individual based ASCVD risk prediction in Chinese is uncertain. Although guidelines generally recommend risk assessment to drive prevention and treatment strategies, adherence to these recommendations are low in clinical practice,^{42,43} especially in primary care settings.⁴³ Furthermore, cost-effectiveness and feasibility of more complex and accurate risk assessment tools for clinical application are unknown. Moreover, no studies to date

have evaluated if ASCVD risk prediction calculators developed from Chinese cohorts can be applied to people living in other East Asian countries or Chinese persons living abroad, including Europe and the United States.

ASCVD RISK PREDICTION IN JAPAN. Since 1960, when the first epidemiologic studies of CVD were performed in Japan, high stroke mortality rates but low CHD mortality rates were observed.⁴⁶ Although stroke mortality has significantly improved in subsequent decades, CHD mortality remains low compared with Western populations. The 2012 Japan Atherosclerosis Society (JAS) Guidelines for Prevention of Atherosclerotic Cardiovascular Diseases was the first guideline to introduce absolute risk assessment of ASCVD and used the NIPPON DATA80 risk chart to predict 10-year CHD mortality.⁴⁴ Similarly, the 2019 Japanese Society of Hypertension incorporated absolute risk assessment as a reference to inform blood pressure treatment thresholds and management.⁴⁵

In the 2017 JAS guideline,⁴⁶ the Suita score was able to accurately estimate the absolute incidence of CHD by incorporating demographics and risk factors including age, sex, smoking, blood pressure level, HDL-C, LDL-C, impaired glucose tolerance, and family history of premature CHD. The Suita score was chosen from 10 different published risk prediction scores in Japan where internal validation was carefully performed. Although attempts were made to validate the predictive model for the guidelines in an external population, even within Japan, cardiovascular disease risk differs between urban and nonurban areas. Thus, it was extremely difficult to set a standard population for external validation.⁴⁷

Based on the algorithm for ASCVD risk assessment, individuals are first screened to determine if they should be categorized as primary or secondary prevention candidates. They are also evaluated for high-risk comorbidities including diabetes, chronic kidney disease, noncardiogenic stroke, and peripheral artery disease. If none of these conditions are present, the Suita score is calculated and individuals are stratified into low-, moderate-, and high-risk categories. The LDL-C targets in primary prevention are set at <160 mg/dL for low risk, <140 mg/dL for moderate risk, and <120 mg/dL for high risk, respectively. The LDL-C target for patients with established CHD is <100 mg/dL, but if they also have a history of familial hypercholesterolemia or acute coronary syndrome, then a lower target LDL-C level of <70 mg/dL should be considered. Patients with both diabetes

TABLE 2 ASCVD Risk Assessment Tools in Japan

First Author	Study Population	Predictors (Other Than Age and Sex)	Outcomes	Duration of Risk Prediction, y	Validation	Risk Tools
Iso et al ⁵⁴	10,334 men and 19,542 women aged 40-69 y	hsCRP SBP, DBP, DM, HDL, LDL, TC, alcohol, BMI, smoking habits, and medications	ASCVD	10	Internal	Odds ratio
Nippon Data80 et al ⁴⁴	4,098 men and 5,255 women age 30 y or higher	SBP, smoking, TC and glucose levels	ASCVD	19	Internal	Risk charts, Cox proportional hazard models
Hirai et al ⁴⁷	14,072 men and 21,307 women Women aged 40-79 y	SBP, DM, HDL, LDL, CKD stage, smoking habits	CHD	3	Internal	Suita risk score
Honda et al ⁸⁴	1,026 men and 1,428 women aged 40-84 y	SBP, DM, HDL, LDL, proteinuria, smoking habits, and regular exercise.	ASCVD	24	Internal	Cox proportional hazard model

CKD = chronic kidney disease; hsCRP = high sensitivity C-reactive protein; other abbreviations as in Table 1.

and ASCVD should also be treated to an aggressive LDL-C goal of <70 mg/dL.

It is important to note that the absolute ASCVD risk estimated by the Suita score only includes CHD and not stroke, unlike the PCE and SCORE2 risk calculators that include both. In Japan, however, cerebral hemorrhage accounts for a high proportion of strokes, whereas the percentages of lacunar, cardioembolic, and atherothrombotic strokes are almost equivalent,^{48,49} with the first 2 not being associated with hypercholesterolemia.⁴⁸ Although Japanese people have a high incidence of stroke, they have a relatively low proportion of stroke phenotypes associated with dyslipidemia, and therefore it has been difficult to use “stroke” when setting lipid management targets. In contrast, hypercholesterolemia is a strong risk factor for atherothrombotic infarction. Accordingly, the new 2022 JAS guideline, uses a recently published risk score from the Hisayama study that predicts incidence of the combined outcome of CHD and atherothrombotic cerebral infarction for individual risk assessment.⁵⁰ A summary of ASCVD calculators used in Japan is shown in Table 2.

The new JAS guideline also reviews the application of imaging and physiological measurements for sub-clinical atherosclerosis detection and their additive value to risk prediction models. These markers include the following: 1) periventricular hyperintensity or deep and subcortical white matter hyperintensity by brain magnetic resonance imaging; 2) carotid intima-media thickness and plaque characteristics; 3) CAC; 4) pulse wave velocity; 5) cardio-ankle vascular index; and 6) ankle-brachial index (ABI). The authors reviewed whether these potential risk enhancers improve the predictive ability of

traditional risk factors for the development of ASCVD in a primary care setting. A meta-analysis of Japanese studies showed that brachial-ankle pulse wave velocity (baPWV) improves risk prediction,⁵¹ but the model was not generated from a Japanese population. Therefore, the applicability to set baPWV cutoff values for Japanese people remains elusive. Other markers, including magnetic resonance imaging, carotid intima-media thickness, and plaque characteristics, and CAC score still lack evidence for risk prediction, and ABI did not enhance risk prediction models.

For high-risk individuals, such as stable patients with suspected CHD, the addition of CAC score to a risk prediction model that employs traditional risk factors significantly improved CHD risk classification.⁵² Given the high availability of computed tomography (CT) scanners in Japan (highest per capital in the world, 115.7 per million people in 2020),⁵³ many studies have been conducted to probe the additive value of coronary CT angiography and CAC score for symptomatic patients. However, use of these imaging modalities for comprehensive risk assessment have not been performed. This is also true for other imaging and physiological measurements, such as magnetic resonance imaging, carotid intima-media thickness, plaque characteristics, baPWV, and cardio-ankle vascular index. Biomarkers such as C-reactive protein,⁵⁴ small dense LDL,⁵⁵ and modified lipoproteins (LOX-1 ligands containing apolipoprotein B,⁵⁶ malondialdehyde-modified LDL⁵⁷) have also been validated in epidemiological and clinical studies in Japan and have been found to refine ASCVD risk when added to classical risk factors. However, they have only been assessed at the individual level, and

TABLE 3 ASCVD Risk Assessment Tools in Korea

First Author	Study Population	Predictors (Other Than Sex and Age)	Outcomes (Follow-Up)	Duration of Risk Prediction, y	Validation	Risk Tools Provided
Jee et al ⁶²	777,502 men and 446,238 women aged 30-84 y	SBP, TC, diabetes, smoking, activity, BMI, alcohol	Stroke	13	Internal	Equation, risk score table
Jee et al ⁶⁸	164,500 men and 104,310 women aged 30-74 y	BP, TC, diabetes, smoking, (HDL-C, LDL-C, TG; optional)	CHD	12	Internal	Equation, HR
Park et al ⁶⁵	3,135 men and 2,425 women aged 43-63 y	Hypertension, diabetes, SBP, DBP, TC, HDL-C, LDL-C, smoking, atrial fibrillation, white blood cells, creatinine, hemoglobin A1c	ASCVD	3	Internal	HR
Jung et al ⁶¹	119,715 men and 80,295 women aged 40-79 y	SBP, TC, HDL-C, DM, smoking, antihypertensive	ASCVD	10	Internal	Equation
Cho et al ⁶³	152,076 men and 144,354 women aged 40-79 y	BMI, SBP, DBP, TC, FBG, smoking, activity	ASCVD	10	Internal and External	Machine learning algorithm
Cho et al ⁶⁵	222,998 individuals aged 40-70 y	SBP, DBP, TC, HDL-C, LDL-C, TG, FBG, GGT, GFR, proteinuria, BMI, WC, smoking, alcohol, activity, antidiabetic, antihypertensive, statins, family history of CHD, family history of stroke	ACS + stroke	8	Internal	HR, nomogram
Choi et al ⁶⁶	10,412,947 men and 11,168,849 women Aged 40-79 y	SBP, TC, HDL-C, diabetes, smoking, antihypertensive, (+8 optional markers)	ASCVD	5	Internal	Machine learning algorithm
(Unpublished)	43,798 individuals aged 40-79 y	SBP, TC, HDL-C, diabetes, smoking, antihypertensive	ASCVD	10	Internal and external	Equation, risk chart

ACS = acute coronary syndrome; DBP = diastolic blood pressure; FBG = fasting blood glucose; GFR = glomerular filtration rate; GGT = gamma-glutamyl transferase; TG = triglycerides; other abbreviations as in Table 1.

without validation in population-based cohorts, are difficult to incorporate into risk prediction models. They have limited clinical application since they are not routinely measured at clinic visits.

Since the establishment of universal health coverage for all Japanese people in 1961, there are limited incentives to use imaging modalities such as CT scans to enhance systematic risk prediction or to evaluate the cost-effectiveness of these tests, especially for high-risk individuals. Future directions should focus on determining whether biomarkers or subclinical atherosclerosis detection can improve risk assessment for low- and intermediate-risk people.

ASCVD RISK PREDICTION IN THE REPUBLIC OF KOREA (SOUTH KOREA). Early Korean studies attempted to recalibrate risk assessment tools developed in Western populations because they overestimated the absolute risk of ASCVD in the Korean population.⁵⁸⁻⁶¹ Later studies used Korean data to develop risk prediction models for stroke, coronary artery disease, or ASCVD.⁶¹⁻⁶⁸ The National Health Insurance Service, which operates a free health screening program for all Korean adults, provides risk

assessment for CVD and diabetes based on individualized data. Private health screening centers and research institutes are also using disease prediction on common chronic diseases, including ASCVD, diabetes, cancer, and Alzheimer’s disease. However, ASCVD risk prediction assessment has not yet been adopted as the primary tool for risk assessment in clinical practice guidelines.⁶⁹ A summary of available ASCVD risk assessment tools in Korea is provided in Table 3.

The Korean Society of Lipid and Atherosclerosis recently released the fifth edition of the Korean Guidelines for the Management of Dyslipidemia. Compared with the fourth edition, the revised guidelines lowered the LDL-C target level for patients with CHD from <70 to <55 mg/dL.^{70,71} The risk classification of people with diabetes was also subdivided based on recent Korean data,⁷² with the LDL-C target value lowered in selected high-risk diabetes groups.⁷³ Lower LDL-C targets are also recommended for patients with significant carotid artery stenosis or abdominal aortic aneurysm. Although risk classification based on the presence of specific risk factors remained, it is noteworthy that the fifth KSOLA

guidelines introduced, for the first time, an optional recommendation to use a specific ASCVD prediction model. The guidelines noted that "ASCVD risk scores based on the Korean Genome and Epidemiology Study (KOGES) cohort can be used as a risk enhancer in low- or moderate-risk groups.⁷¹ The reason why it is recommended to be used only in low- to moderate-risk groups is that the absolute incidence of ASCVD is very low in Koreans, so if they rely only on the ASCVD prediction model, some patients high-risk groups with major modifiable risk factors may not be identified for treatment.⁷¹ The Korean Society of Hypertension (KSH) released the 2022 KSH Guidelines for the Management of Hypertension without significant changes in risk classification compared with the previous 2018 KSH guideline. Assessment of CVD risk and blood pressure targets are still based on the presence or absence of clinical ASCVD or the number of established cardiovascular risk factors. The KSH guidelines acknowledge the importance of accurate individual risk assessment for selecting the best ASCVD prevention strategy and recommend the use of risk assessment tools but does not endorse a specific risk prediction model.^{71,74}

There are several reasons why ASCVD risk prediction is mainly of academic interest and is not widely accepted in clinical practice guidelines in Korea. Because most of the ASCVD risk prediction models developed in Korea are based on data from limited health screening centers, there is criticism about the lack of adequate representation. Additionally, the ASCVD risk prediction models were mostly validated internally without external validation in a diverse cohort population. Furthermore, there is insufficient evidence demonstrating the clinical efficacy and cost-effectiveness of preventive pharmacological interventions. These deficiencies make it difficult to recommend an ASCVD risk prediction model in daily clinical practice.⁶⁹ Risk factors in the Korean ASCVD risk prediction models are similar to those used in other countries. The major risk predictors include sex, age, blood pressure, smoking, diabetes, and TC. It is notable that many of the Korean prediction models only use TC instead of LDL-C and HDL-C because Korea's general national health screening program often only measures TC levels.^{61-63,68} In addition, the incidence of CHD is low, and LDL-C and TC levels are strongly correlated to risk in the Korean population while the predictive power of TC is not inferior to LDL-C plus HDL-C levels.^{56,66} Efforts to use calculators, including the SCORE2 (Systemic Coronary Risk Evaluation 2) and SCORE2-OP (SCORE2-Older persons), resulted in an underestimation risk in young Korean men and women (aged 40-59 years)

and overestimation in older individuals, highlighting the need for a region-specific calculator.⁷³

The most recently developed Korean ASCVD prediction model is based on data from over 150,000 participants in the Korean Genome and Epidemiology Study. The risk predictors are similar to previous models, but this model includes large-scale prospective cohort data and was validated both internally and externally. In addition, the new model may have better clinical utility because the incidence of CVD in Koreans is rapidly increasing. Although ASCVD risk scores are not officially adopted in clinical practice guidelines, many CVD risk prediction studies have been published in Korea utilizing new biomarkers and imaging modalities.⁷⁵⁻⁸¹ Several of them have demonstrated improvements in ASCVD predictive power using CAC or CTA and employing machine learning methods. Earlier studies cross-sectionally compared CAC and Western-derived ASCVD risk scores among asymptomatic individuals and reported that high CAC levels are also observed among some Korean persons with low to moderate ASCVD risk.^{75,77} These findings indirectly support the usefulness of CAC-based risk classification. Later studies added more direct evidence of CAC score in the risk classification of asymptomatic Korean adults. The CONFIRM (COronary CT Angiography Evaluation For Clinical Outcomes InteRnational Multicenter) study evaluated the clinical utility of CAC and coronary CTA in Korean as well as European/American populations.⁷⁴ The CONFIRM study reported that CAC improved risk stratification and provided incremental value beyond FRS for predicting major adverse cardiac events.⁷⁴ Another Korean study reported that the addition of CAC to FRS improved risk prediction for CVD mortality in young adults but not in older individuals.⁸² The predictive value of coronary CTA was also evaluated in combination with CAC. In a 2-year follow-up study of Koreans participating in health screens, both CAC and coronary artery stenosis on coronary CTA are independent predictors of CVD outcomes. However, when CAC and coronary CTA stenosis were evaluated simultaneously, CVD risk was associated only with stenosis on coronary CTA but not with CACS.⁷⁶ In the CONFIRM study, coronary CTA provides incremental predictive value for asymptomatic individuals with moderate CAC scores (100-400), but not for lower or higher CAC scores.⁷⁸ There are also studies that evaluated the usefulness of funduscopy.^{80,81} A Korean study found that virtual assessment of CAC estimated from deep learning analysis of retinal photographs is comparable to CT-measured CAC in predicting CVD events and improves current risk stratification for cardiovascular

events.⁸¹ If further validated, funduscopy plus deep learning algorithms have the potential to serve as a cost-effective and radiation-free alternative to measuring CAC, particularly in resource-limited settings. Collectively, these studies suggest that the inclusion of imaging strategies to improve the accuracy of risk calculators such as SCORE2 and SCORE2-OP in young and old Koreans, respectively.⁸³

FUTURE DIRECTIONS AND CONCLUSIONS. In this overview of ASCVD risk assessment in East Asian countries, specifically China, Japan, and South Korea, ASCVD risk is significantly overestimated, in particular CHD, when applying calculators developed in the United States including the FRS and PCE. Unlike Europe and the United States, incidence of CHD is much lower while stroke rates are higher in Japan, Korea, and China. Studies to recalibrate these risk scores have been unsatisfactory, resulting in each country developing their own risk prediction scores based on epidemiologic studies using native cohorts and traditional risk factors.^{29,84,85} Unfortunately, many of these national risk prediction scores lack external validation and generalizability. Furthermore, there are challenges to implementation and adherence of guidelines in clinical practice. Application of additional risk enhancers such as CAC score and biomarkers have not been extensively studied to determine their ability to improve risk prediction models. Moreover, the utility of using specific biomarkers in risk prediction algorithms may be limited if they are not routinely measured in clinical practice. Each country's risk score was developed using Western risk calculators as a framework, and no cross validation has been performed between among China, Japan, and South Korea, nor among a larger aggregate that includes other East Asian countries. This may provide an opportunity to develop a more refined regional risk score and new risk indices through collaboration given similarities in ASCVD prevalence, disease characteristics, and lifestyles.

Although the current landscape of ASCVD risk prediction tools in East Asian countries is improving, there is still significant need for refinement. External validation, implementation, and adherence are significant challenges. Consequently, there is a great need for multinational approaches for the conduct of registries and clinical trials in East Asian countries and beyond. Region-specific standardized protocols for risk factor assessment and ASCVD outcomes should be created to improve generalizability of these risk prediction models. As discussed previously, although United States-based scoring algorithms can discriminate ASCVD risk reasonably well in Asian

populations, when recalibrated,⁸⁶ they substantially overestimate actual risk in Asian cohorts. The Asia Pacific Cohort Studies Collaboration (APCSC) risk equation for predicting 8-year risk of a major cardiovascular event was the first unified attempt to develop such a model, but only involved China, Japan, Korea, and Singapore,⁸⁷ and having been conducted 2 decades ago, is in need of further updating in a larger set of Asian-Pacific countries.

The new risk prediction models being developed in East Asian countries should standardize the definition of ASCVD and should include CHD, stroke, and PAD. Moreover, future calculators should also consider prediction of other CVD conditions, notably heart failure and atrial fibrillation given their important burden among CVD conditions. Given the higher rates of cerebrovascular disease relative to CHD in the region, standardization of these risk calculators may promote better opportunities for cross-validation. As highlighted in the previous text, risk-enhancing factors have been validated in Western populations but may perform differently in Asian cohorts. Therefore, identification of risk-enhancing factors unique to East Asian populations should be characterized to improve risk classification. Because they are easy to perform and easily understood by patients, clinical implementation of race- and ethnic-specific ASCVD risk calculators should not be difficult and welcomed by providers who yearn for more accurate risk estimators for patient subgroups. Although they may not be readily available, established risk modifiers such as subclinical atherosclerosis detected by noninvasive imaging such as CAC, carotid ultrasound, and ABI should be studied more extensively in East Asian countries and may improve risk refinement with the caveat that many patients receiving these imaging studies may be at higher risk for ASCVD, which results in selection bias and may affect the accuracy of risk assessment. At the 2023 European Society of Cardiology meeting, a new regional risk prediction model for Asia, SCORE2-ASIA, was revealed. This new risk assessment tool for Asia will be recalibrated using multiplication factors based on average CV risk in each Asian region (categorized as low-moderate, high, or very high risk). Once these risk prediction models are both internally and externally validated, they could be tested in Asian American cohorts.

The challenge to validating these risk models in U.S. cohorts is the lack of participation of East Asian Americans in clinical trials and their low representation in national health surveys. More importantly, disaggregation of Asian Americans participating in clinical trials and surveys is lacking. Although the NHIS provides disaggregated Asian-American

HIGHLIGHTS

- Prior studies have shown that the prevalence of ASCVD and its risk factors vary across East Asian countries, but whether these differences extend to East Asians living in the United States and how they affect the validity of ASCVD-risk calculators in East Asians in the United States and in their native countries remain unclear.
- ASCVD risk calculators, developed by the ACC/AHA, overestimate risk in Chinese, Koreans, and Japanese people. Risk calculators developed by professional organizations in the countries of origin, however, lack validation in East Asians living in the United States, who are exposed to different environmental and cultural influences.
- Future studies should include individuals from East Asian subpopulations and report disaggregated findings to inform the development of an ASCVD-risk calculator specific to East Asian U.S. immigrants that can then be prospectively validated.

subgroups, this data is only available for 3 of the largest Asian subgroups (Asian Indian, Chinese, and Filipino persons) because of the small number of participants for other Asian subgroups (mostly Japanese, Koreans, and Vietnamese persons). Although the NHANES oversampled for Asian American participants for several years, access to disaggregated Asian subgroup data is restricted. While East Asian countries represent more than one-quarter of the world's population, representation of subjects in global clinical trials that inform international practice guidelines is far less. In the REDUCE-IT (Reduction of Cardiovascular Events With Icosapent Ethyl-Intervention Trial) trial of icosapent ethyl,⁸⁸ only 3% of patients were from the Asia-Pacific region while 12% of participants were from Asia in the ODYSSEY OUTCOMES (Evaluation of Cardiovascular Outcomes After an Acute Coronary Syndrome During Treatment With Alirocumab) trial of alirocumab.⁸⁹ One spotlight is the American College of Cardiology NCDR (National Cardiovascular Data Registry) that recently began to include disaggregated Asian and Hispanic subgroups for many of its registries.

Until there is a critical mass of Asian American participants in clinical trials and registries, it will be exceedingly difficult to create a more refined ASCVD risk assessment tool for East Asian Americans or other Asian American subgroups such as South Asian Americans. A call to action through advocacy and federal legislation that requires health care systems and national health agencies, including the Centers for Disease Control and Prevention, to collect and report disaggregated Asian American subgroup information may be necessary. Recently, the National Heart, Lung, and Blood Institute announced a new epidemiological cohort study to address key population health gaps for Asian Americans, Native Hawaiians, and Pacific Islanders.⁹⁰ The 7-year study (MOSAAIC [Multi-ethnic Observational Study in American Asian and Pacific Islander Communities]) was launched in August 2023 and will recruit 10,000 adults, ages 18 to 64 years, from 5 U.S. sites. It will focus on cardiovascular health as well as other conditions including lung health, mental health, and social determinants of health in individuals who self-identify as having ancestral background from East Asia, South Asia, or Southeast Asia; or who self-identify as Native Hawaiian and/or Pacific Islander. This study is a major step forward to enhance research efforts of Asian subgroups in the United States and may help to accelerate initiation of studies for these populations.

ACKNOWLEDGEMENT The authors would like to thank Colin Lai for his contribution to formatting the figures in this review.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

Dr Nguyen has received support from the American Heart Association (AHA 23SCISA1145819). Dr Okamura has received funding support from Grant-in-Aid for Scientific Research (A) (21H04854) and Health and Labour Sciences Research Grants, Japan (Comprehensive research on lifestyle related diseases including cardiovascular diseases and diabetes mellitus) (24FA1002). Dr Wong has received funding through his institution from Amgen, Novartis, Novo Nordisk and Regeneron; and is a consultant for Amgen, Heart Lung, Ionis, and Novartis. Dr Yang is supported by the UW Medicine Asian Health Initiative and the Carl and Renée Behnke Endowed Chair for Asian Health; has received research funding from Microsoft Research; is a stockholder and advisory board member for Measure Labs; and is a consultant for Idorsia, Genentech, and Mineralys. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

ADDRESS FOR CORRESPONDENCE: Dr Eugene Yang, University of Washington School of Medicine, 1959 NE Pacific Street, Box 356005, Seattle, Washington 98195, USA. E-mail: eyang01@uw.edu. X handle: [@DrEugeneYang](https://twitter.com/DrEugeneYang).

REFERENCES

1. United Nations: Population Division Data Portal Accessed June 2024. <https://population.un.org/dataportal/data/indicators/49/locations/>
2. Zhao D. Epidemiological features of cardiovascular disease in Asia. *JACC Asia*. 2021;1(1):1-13. <https://doi.org/10.1016/j.jacasi.2021.04.007>
3. O'Keefe PB, Umapathi N, Somanathan A, et al. Live long and prosper: aging in East Asia and Pacific (English). Washington, D.C.: World Bank Group. Accessed June 1, 2024. <http://documents.worldbank.org/curated/en/832271468184782307>
4. U.S. Census Bureau. 2019 American Community Survey. Accessed March 1, 2021. <https://www.census.gov/programs-surveys/acs/data.html>
5. Volgman AS, Palaniappan LS, Aggarwal NT, et al. Atherosclerotic cardiovascular disease in South Asians in the United States: Epidemiology, risk factors, and treatments: A Scientific Statement From the American Heart Association. *Circulation*. 2018;138(1):e1-e34.
6. DeFilippis AP, Young R, Carrubba CJ, et al. An analysis of calibration and discrimination among multiple cardiovascular risk scores in a modern multiethnic cohort. *Ann Intern Med*. 2015;162(4):266-275.
7. Patel AP, Wang M, Kartoun U, Ng K, Khera AV. Quantifying and understanding the higher risk of atherosclerotic cardiovascular disease among South Asian individuals: Results from the UK Biobank Prospective Cohort Study. *Circulation*. 2021;144(6):410-422.
8. Koirala B, Turkson-Ocran RA, Baptiste D, et al. Heterogeneity of cardiovascular disease risk factors among Asian Immigrants: insights from the 2010 to 2018 National Health Interview Survey. *J Am Heart Assoc*. 2021;10(13):e020408.
9. Khan SS, Matsushita K, Sang Y, et al. Development and validation of the American Heart Association's PREVENT Equations. *Circulation*. 2024;149(6):430-449.
10. Hageman SHJ, Huang Z, Lee H, et al. SCORE2 Asia-Pacific Writing Group, the European Society of Cardiology and European Association of Preventive Cardiology: Cardiovascular Risk Collaboration (ESC CRC), the ASEAN Federation of Cardiology (AFC), the Asian-Pacific Society of Cardiology (APSC), on behalf of the SCORE2 Asia-Pacific Collaborators. Risk prediction of cardiovascular disease in the Asia-Pacific region: the SCORE2 Asia-Pacific model. *Eur Heart J*. Published online September 1, 2024. <https://doi.org/10.1093/eurheartj/ehae609>
11. *Countries of Asia: East Asia*. Nations Online; 2021. Accessed March 2024. <https://www.nationsonline.org/oneworld/asia.htm#East-Asia>
12. Summary Health Statistics. *National Health Interview Survey*. 2018. Accessed June 1, 2024. https://ftp.cdc.gov/pub/Health_Statistics/NCHS/NHIS/SHS/2018_SHS_Table_A-18.pdf
13. Xu J, Murphy S, Kochanek K, Arias E. National vital statistics report - deaths: final data for 2019-2021; p. 1-87. Accessed June 1, 2024. <https://www.cdc.gov/nchs/data/nvsr/nvsr70/nvsr70-08-508.pdf>
14. Holland AT, Wong EC, Lauderdale DS, Palaniappan LP. Spectrum of cardiovascular diseases in Asian-American racial/ethnic subgroups. *Ann Epidemiol*. 2011;21(8):608-614.
15. Pursnani S, Merchant M. South Asian ethnicity as a risk factor for coronary heart disease. *Atherosclerosis*. 2020;315:126-130.
16. Hastings KG, Jose PO, Kapphahn KI, et al. Leading causes of death among Asian American subgroups (2003-2011). *PLoS One*. 2015;10(4):e0124341.
17. Jose PO, Frank AT, Kapphahn KI, et al. Cardiovascular disease mortality in Asian Americans. *J Am Coll Cardiol*. 2014;64(23):2486-2494.
18. Shah NS, Xi K, Kapphahn KI, et al. Cardiovascular and cerebrovascular disease mortality in Asian American subgroups. *Circ Cardiovasc Qual Outcomes*. 2022;15(5):e008651.
19. NCD-RISC: NCD Risk Factor Collaboration. Accessed March 2024. <https://www.ncdrisc.org/>
20. Commodore-Mensah Y, Selvin E, Aboagye J, et al. Hypertension, overweight/obesity, and diabetes among immigrants in the United States: an analysis of the 2010-2016 National Health Interview Survey. *BMC Public Health*. 2018;18(1):773.
21. Chiu JF, Bell AD, Herman RJ, et al. Cardiovascular risk profiles and outcomes of Chinese living inside and outside China. *Eur J Cardiovasc Prev Rehabil*. 2010;17(6):668-675.
22. Hirooka N, Takedai T, D'Amico F. Assessing physical activity in daily life, exercise, and sedentary behavior among Japanese moving to westernized environment: a cross-sectional study of Japanese migrants at an urban primary care center in Pittsburgh. *Asia Pac Fam Med*. 2014;13(1):3.
23. Li Y, Zhu A, LE A, et al. Association of acculturation with cardiovascular risk factors in Asian-American subgroups. *Am J Prev Cardiol*. 2023;13:100437.
24. Grundy SM, Stone NJ, Bailey AL, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on clinical practice guidelines. *J Am Coll Cardiol*. 2019;73(24):3168-3209.
25. Liu J, Hong YL, D'Agostino RB, et al. Predictive value for the Chinese population of the Framingham CHD risk assessment tool compared with the Chinese multi-provincial cohort study. *JAMA*. 2004;291(21):2591-2599.
26. Wu Y, Liu X, Li X, et al. Estimation of 10-year risk of fatal and nonfatal ischemic cardiovascular diseases in Chinese adults. *Circulation*. 2006;114(21):2217-2225.
27. Yang X, Li J, Hu D, et al. Predicting the 10-year risks of atherosclerotic cardiovascular disease in Chinese population: The China-PAR Project (Prediction for ASCVD Risk in China). *Circulation*. 2016;134(19):1430-1440.
28. Wang W, Zhao D, Liu J. Prospective study on the predictive model of cardiovascular disease risk in a Chinese population aged 35-64. *Chin J Cardiol*. 2003;31:902-908.
29. Wang M, Jing L, Dong Z. A new protocol for risk assessment of atherosclerotic cardiovascular disease in China. *Chin J Cardiol*. 2018;46:87-91. <https://doi.org/10.3760/cma.j.issn.0253-3758.2018.02.003>
30. Wang Y, Liu J, Wang W, et al. Lifetime risk for cardiovascular disease in a Chinese population: the Chinese multi-provincial cohort study. *Eur J Prev Cardiol*. 2020;22(3):380-388.
31. Xing X, Yang X, Liu F, et al. Predicting 10-year and lifetime stroke risk in Chinese population. *Stroke*. 2019;50(9):2371-2378.
32. Writing Committee of Chinese Hypertension Guidelines. 1999 Chinese guidelines for the management of hypertension. *Chin J Hypertens*. 2000;94-102.
33. Joint Committee for Developing Chinese Guidelines on Prevention and Treatment of Dyslipidemia in Adults. Chinese guidelines on prevention and treatment of dyslipidemia in adults. *Chin J Cardiol*. 2007;35:390-416.
34. Chinese Society of Cardiology of Chinese Medical Association. Guideline on the primary prevention of cardiovascular diseases. *Cardiol Discovery*. 2021;1:70-104.
35. Joint Committee for Guideline Revision. 2016 Chinese guidelines for the management of dyslipidemia in adults. *J Geriatr Cardiol*. 2018;15:1-29.
36. Task Force on Chinese Guidelines for the Prevention of Cardiovascular Diseases. Chinese guidelines for the prevention of cardiovascular diseases. *Chin J Cardiol*. 2018;46:10-25.
37. Hou ZH, Lu B, Gao Y, et al. Prognostic value of coronary CT angiography and calcium score for major adverse cardiac events in outpatients. *JACC Cardiovasc Imaging*. 2012;5:990-999.
38. Detrano R, Guerci AD, Carr JJ, et al. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. *N Engl J Med*. 2008;358:1336-1345.
39. Atherosclerosis and Coronary Heart Disease Working Group of Chinese Society of Cardiology. Chinese expert consensus on lipid management of very high-risk atherosclerotic cardiovascular disease patients. *Chin J Cardiol*. 2020;48:280-286.
40. Liu X, Shen P, Zhang D, et al. Evaluation of atherosclerotic cardiovascular risk prediction models in China. *JACC: Asia*. 2022;2:33-43.
41. Qi Y, Fan J, Liu J, et al. Cholesterol-overloaded HDL particles are independently associated with progression of carotid atherosclerosis in a cardiovascular disease-free population: a community-based cohort study. *J Am Coll Cardiol*. 2015;65(4):355-363.
42. Shang P, Liu GG, Zheng X, et al. Association between medication adherence and 1-Year major cardiovascular adverse events after acute myocardial infarction in China. *J Am Heart Assoc*. 2019;8(9):e011793.
43. Lu Y, Zhang H, Lu J, et al. Prevalence of dyslipidemia and availability of lipid-lowering

- medications among primary health care settings in China. *JAMA Netw Open*. 2021;4(9):e2127573.
44. Nippon Data Research Group. Risk assessment chart for death from cardiovascular disease based on a 19-year follow-up study of a Japanese representative population NIPPON DATA80. *Circ J*. 2006;70:1249-1255.
45. Umemura S, Arima H, Arima S, et al. The Japanese Society of Hypertension Guidelines for the Management of Hypertension (JSH 2019). *Hypertens Res*. 2019;42(9):1235-1481.
46. Kinoshita M, Yokote K, Arai H, et al. Japan Atherosclerosis Society (JAS) Guidelines for prevention of atherosclerotic cardiovascular diseases 2017. *J Atheroscler Thromb*. 2018;25(9):846-984.
47. Hirai H, Asahi K, Yamaguchi S, et al. New risk prediction model of coronary heart disease in participants with and without diabetes: assessments of the Framingham risk and Suita scores in 3-year longitudinal database in a Japanese population. *Sci Rep*. 2019;9(1):2813.
48. Imamura T, Doi Y, Arima H, et al. LDL cholesterol and the development of stroke subtypes and coronary heart disease in a general Japanese population: the Hisayama study. *Stroke*. 2009;40(2):382-388.
49. Toyoda K, Yoshimura S, Nakai M, et al. Twenty-year change in severity and outcome of ischemic and hemorrhagic strokes. *JAMA Neurol*. 2022;79(1):61-69.
50. Okamura T, Tsukamoto K, Arai H, et al. Japan Atherosclerosis Society (JAS) guidelines for prevention of atherosclerotic cardiovascular diseases 2022. *J Atheroscler Thromb*. 2024;31:641-853. s.
51. Ohkuma T, Ninomiya T, Tomiyama H, et al. Brachial-ankle pulse wave velocity and the risk prediction of cardiovascular disease: An individual participant data meta-analysis. *Hypertension*. 2017;69(6):1045-1052.
52. Nakao YM, Miyamoto Y, Higashi M, et al. Sex differences in impact of coronary artery calcification to predict coronary artery disease. *Heart*. 2018;104:1118-1124.
53. OECD. OECD data explorer. Accessed March 2024. https://stats.oecd.org/index.aspx?DataSetCode=HEALTH_STAT
54. Iso H, Noda H, Ikeda A, et al. The impact of C-reactive protein on risk of stroke, stroke subtypes, and ischemic heart disease in middle-aged Japanese: the Japan Public Health Center-Based Study. *J Atheroscler Thromb*. 2012;19:756-766.6.
55. Higashioka M, Sakata S, Honda T, et al. The association of small dense low-density lipoprotein cholesterol and coronary heart disease in subjects at high cardiovascular risk. *J Atheroscler Thromb*. 2021;28:79-89.
56. Inoue N, Okamura T, Kokubo Y, et al. LOX Index, a novel predictive biochemical marker for coronary heart disease and stroke. *Clin Chem*. 2010;56:550-587.
57. Arai K, Koba S, Yokota Y, et al. Relationships of fatty acids, delta-5 desaturase activity, and lipid profiles in men with acute coronary syndrome. *J Atheroscler Thromb*. 2020;27(11):1216-1229.
58. Liu J, Hong Y, D'Agostino RB, et al. Predictive value for the Chinese population of the Framingham CHD risk assessment tool compared with the Chinese Multi-Provincial Cohort Study. *JAMA*. 2004;291(21):2591-2599.
59. Ahn KA, Yun JE, Cho ER, Nam CM, Jang Y, Jee SH. Framingham equation model overestimates risk of ischemic heart disease in Korean men and women. *Kor J Epidemiol*. 2006;28(2):162-170.
60. Barzi F, Patel A, Gu D, et al. Cardiovascular risk prediction tools for populations in Asia. *J Epidemiol Community Health*. 2007;61(2):115-121.
61. Jung KJ, Jang Y, Oh DJ, et al. The ACC/AHA 2013 pooled cohort equations compared to a Korean Risk Prediction Model for atherosclerotic cardiovascular disease. *Atherosclerosis*. 2015;242(1):367-375.
62. Jee SH, Park JW, Lee SY, et al. Stroke risk prediction model: a risk profile from the Korean study. *Atherosclerosis*. 2008;197(1):318-325.
63. Cho I-J, Sung JM, Chang H-J, Chung N, Kim HC. Incremental value of repeated risk factor measurements for cardiovascular disease prediction in middle-aged Korean adults: results from the NHIS-HEALS (National Health Insurance System-National Health Screening Cohort). *Circ Cardiovasc Qual Outcomes*. 2017;10(11):e004197.
64. Cho IJ, Sung JM, Kim HC, et al. Development and external validation of a deep learning algorithm for prognostication of cardiovascular outcomes. *Kor Circ J*. 2020;50(1):72-84.
65. Cho SY, Kim SH, Kang SH, et al. Pre-existing and machine learning-based models for cardiovascular risk prediction. *Sci Rep*. 2021;11(1).
66. Choi SH, Lee SM, Kim SH, Park M, Yoon HJ. Prediction of 8-year risk of cardiovascular diseases in Korean adult population. *Scientific Reports*. Jul 12 2021;11(1):8886.
67. Park S, Kim YG, Ann SH, et al. Prediction of the 10-year risk of atherosclerotic cardiovascular disease in the Korean population. *Epidemiol Health*. 2023;45:e2023052.
68. Jee SH, Jang Y, Oh DJ, et al. A coronary heart disease prediction model: the Korean Heart Study. *BMJ Open*. 2014;4(5):e005025. <https://doi.org/10.1136/bmjopen-2014-005025>
69. Oh SM, Stefani KM, Kim HC. Development and application of chronic disease risk prediction models. *Yonsei Med J*. 2014;55(4):853-860.
70. Rhee E-J, Kim HC, Kim JH, et al. 2018 Guidelines for the management of dyslipidemia in Korea. *J Lipid Atheroscler*. 2019;8(2):78-131.
71. Kim HL, Lee EM, Ahn SY, et al. The 2022 focused update of the 2018 Korean Hypertension Society Guidelines for the management of hypertension. *Clin Hypertens*. 2023;29:11. <https://doi.org/10.1186/s40885-023-00234>
72. Kim JOR, Jeong Y-S, Kim JH, Lee J-W, Park D, Kim H-S. Machine learning-based cardiovascular disease prediction model: a cohort study on the Korean national health insurance service health screening database. *Diagnostics*. 2021;11(6):943.
73. Yang YS, Kim HL, Kim SH, Moon MK. Lipid management in Korean people with Type 2 diabetes mellitus: Korean Diabetes Association and Korean Society of Lipid and Atherosclerosis consensus statement. *Diabetes Metab J*. 2023;47(1):1-9.
74. Lee JH, Rizvi A, Hartaigh B, et al. The Predictive value of coronary artery calcium scoring for major adverse cardiac events according to renal function (from the coronary computed tomography angiography evaluation for clinical outcomes: An International Multicenter [CONFIRM] Registry). *Am J Cardiol*. 2019;123(9):1435-1442.
75. Sung J, Lim SJ, Choe Y, et al. Comparison of the coronary calcium score with the estimated coronary risk. *Coron Artery Dis*. 2008;19(7):475-479.
76. Park HE, Chun EJ, Choi SI, et al. Clinical and imaging parameters to predict cardiovascular outcome in asymptomatic subjects. *Int J Cardiovasc Imaging*. 2013;29(7):1595-1602. <https://doi.org/10.1007/s10554-013-0235-5>
77. Kim BJ, Kim BS, Kang JH. Conventional vs image-based cardiovascular risk assessment in Korean adults. *Coron Artery Dis*. 2014;25(2):118-124.
78. Cho I, Chang HJ, Oh B, et al. Incremental prognostic utility of coronary CT angiography for asymptomatic patients based upon extent and severity of coronary artery calcium: results from the COronary CT angiography evaluationN For clinical outcomes InteRnational Multicenter (CONFIRM) study. *Eur Heart J*. 2015;36(8):501-508.
79. Lee HY, Shin J, Kim GH, et al. 2018 Korean Society of Hypertension guidelines for the management of hypertension: part II-diagnosis and treatment of hypertension. *Clin Hypertens*. 2019;25:20.
80. Son J, Shin JY, Chun EJ, Jung KH, Park KH, Park SJ. Predicting high coronary artery calcium score from retinal fundus images with deep learning algorithms. *Transl Vis Sci Technol*. 2020;9(6):28. d.
81. Rim TH, Lee CJ, Tham YC, et al. Deep-learning-based cardiovascular risk stratification using coronary artery calcium scores predicted from retinal photographs. *Lancet Digit Health*. 2021;3(5):e306-e316.
82. Kang J, Kim S, Chang Y, Kim Y, Jung HS, Ryu S. Age-stratified effects of coronary artery calcification on cardiovascular and non-cardiovascular mortality in Korean adults. *Eur Heart J Cardiovasc Imaging*. 2022;24(1):17-24.
83. Choi J, Sung S, Park SK, et al. SCORE and SCORE2 in East Asian Population: A Performance Comparison. *JACC Asia*. 2023;4(4):265-274.
84. Honda T, Chen S, Hata J, et al. Development and validation of a risk prediction model for

atherosclerotic cardiovascular disease in Japanese Adults: The Hisayama Study. *J Atheroscler Thromb*. 2022;29(3):345-361. <https://doi.org/10.5551/jat.61960>

85. Park GM, Han S, Kim SH, et al. Model for assessing cardiovascular risk in a Korean population. *Circ Cardiovasc Qual Outcomes*. 2014;7(6):944-951. <https://doi.org/10.1161/circoutcomes.114.001305>

86. Zhang Y, Miao H, Chia Y-C, et al. Cardiovascular risk assessment tools in Asia. *J Clin Hypertens*. 2022;24(4):369-377.

87. Asia Pacific Cohort Studies C, Barzi F, Patel A, et al. Cardiovascular risk prediction tools for populations in Asia. *J Epidemiol Community Health*. 2007;61(2):115-121. <https://doi.org/10.1136/jech.2005.044842>

88. Bhatt DL, Steg PG, Miller M, et al. Cardiovascular risk reduction with icosapent ethyl for hypertriglyceridemia. *N Engl J Med*. 2019;380(1):11-22.

89. Schwartz GG, Steg PG, Szarek M, et al. Alirocumab and cardiovascular outcomes after acute coronary syndrome. *N Engl J Med*. 2018;379(22):2097-2107.

90. Department of Health and Human Services. New epidemiological cohort study among Asian Americans, Native Hawaiians, and Pacific Islanders (AsA-NHPI): Coordinating Center (U24 - Clinical Trial Not Allowed). Accessed August 7, 2022. <https://grants.nih.gov/grants/guide/rfa-files/RFA-HL-23-016.html>

KEY WORDS Asian immigrants, atherosclerosis, China, coronary artery disease, East Asia, Japan, Korea, stroke