Med

Deep Learning-Enabled Screening of Chronic Kidney Disease from Echocardiography

--Manuscript Draft--

Manuscript Number:	MED-D-26-00156R5
Full Title:	Deep Learning-Enabled Screening of Chronic Kidney Disease from Echocardiography
Article Type:	Article
Keywords:	artificial intelligence; deep learning; chronic kidney disease; echocardiography; precision diagnosis
Corresponding Author:	David Ouyang Cedars-Sinai Smidt Heart Institute Pleasanton, UNITED STATES OF AMERICA
First Author:	Victoria Yuan, BS
Order of Authors:	Victoria Yuan, BS
	Hirotaka Ieki
	Alexander Sandhu
	Long H. Nguyen
	Paul P. Cheng
	Stephanie T. Chang
	Andrew P. Ambrosy
	Alan C. Kwan
	Alan S. Go
	Susan Cheng
	David Ouyang
Abstract:	Chronic kidney disease (CKD) affects nearly 850 million individuals globally; the prevalence of undiagnosed CKD is 60%. Taking advantage of the relationship between CKD and cardiovascular disease, we developed a deep learning (DL) model to detect CKD from parasternal long-axis (PLAX) videos using 325,377 PLAX videos from 62,818 patients at Cedars-Sinai Medical Center (CSMC). We externally validated our model in two independent cohorts of 2,224 patients at Stanford Healthcare (SHC) and 41,611 patients at Kaiser-Permanente Northern California (KPNC). In a held-out test cohort at CSMC, our model detected any stage of CKD with an area under the curve (AUC) of 0.756 [95% confidence interval 0.749 – 0.763], with consistently strong performance in KPNC (AUC 0.718 [0.714 – 0.723]) and SHC (AUC 0.719 [0.704 – 0.735]). Our DL echo model detected CKD with robust performance at two external clinical sites, offering an avenue for noninvasive screening and improved detection rates.
Suggested Reviewers:	Maxime Sermesant maxime.sermesant@inria.fr Research expertise in deep learning applied to cardiac imaging
	Azra Bihorac abihorac@ufl.edu Expertise in clinical informatics and nephrology
	Jennifer Ho Beth Israel Deaconess Medical Center jho@bidmc.harvard.edu Expertise in machine learning, cardiology, and cardiovascular-kidney-metabolic syndrome
Opposed Reviewers:	

Additional Information:	
Question	Response
Standardized datasets A list of datatypes considered standardized under Cell Press policy is available here. Does this manuscript report new standardized datasets?	No
Original code Does this manuscript report original computer code, algorithms, or computational models?	Yes
Reviewers must have free and anonymous access to original code, algorithms, and computational models that underlie this work. Please provide a code location and instructions for access below. Without access information, your manuscript cannot be sent to reviewers. Access must not require that users provide any personal information, including personal logins or passwords. Please consult this guide for more information: "How standardized datasets and original code accompany Cell Press manuscripts from submission through publication." Email us at med@cell.com if you have any questions. as follow-up to "**Original code** Does this manuscript report original computer code, algorithms, or computational models? "	Our model weights and code are publicly available at https://github.com/echonet/ckd

25 July 2026

Dear Dr. Le,

Thank you for your thoughtful review of our manuscript. We are pleased to submit a revised version of our manuscript entitled "**Deep Learning-Enabled Screening of Chronic Kidney Disease from Echocardiography**" for consideration for publication in *Med* as an Article.

In this work, we developed a deep learning (DL) model that detects any stage of CKD from parasternal long axis (PLAX) echocardiography. Our model demonstrated strong performance across three geographically distinct clinical centers, with robust performance to patient demographics, co-morbidities, and disease severity at all three sites. We believe that our work can contribute to earlier detection of CKD by offering a rapid, noninvasive screening method. Identifying patients can in turn connect them to more timely therapies. Our model code and weights are publicly available to facilitate its use and support transparency and reproducibility in research.

In this revision, we end our Discussion with a Limitations of Study section, including a statement on the need for testing further thresholds for sensitivity and specificity. Additionally, we added sections on Participant Information and Resource Availability as well as revised typographical errors in the Supplementary Materials. We have also reformatted our manuscript per *Med* guidelines; updated the Author Contributions; increased the brightness in Figure 1 per prior reviewer request; and provided a Graphical Abstract, Highlights, Context and Significance, and eTOC. Our manuscript word count has been updated to exclude the STAR Methods. Our running title is "Deep Learning Screens for CKD on Echocardiography."

We appreciate the thoughtful feedback reviewers and hope it is now acceptable for publication.

Sincerely,

David Ouyang, MD

REPLY: Dear Editors, thank you for your thoughtful and detailed review of our manuscript. In this review, we respond to reviewer feedback in showing that:

1. Our deep learning (DL) model learns CKD-specific signals and these signals persistent even when controlling for cardiometabolic confounders

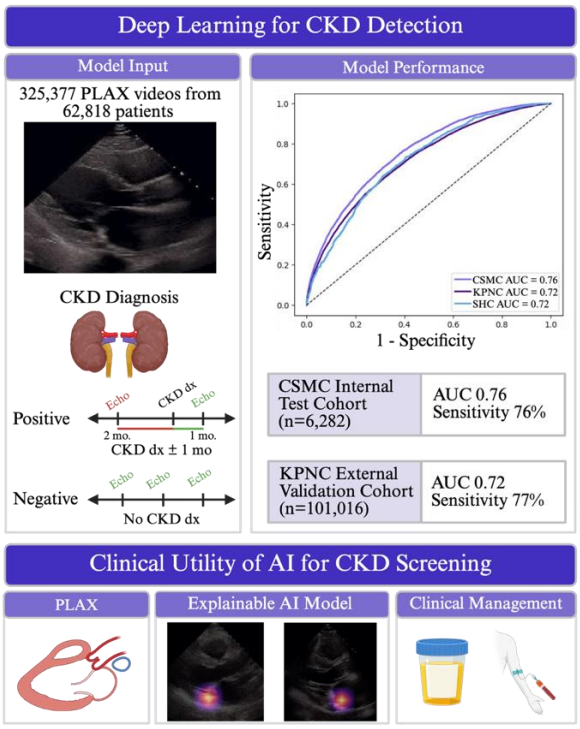
We have modified our **Results: Model Performance** to:

> *We also tested our model in balanced subgroups based on diabetes, hypertension, dyslipidemia, CAD, age ≥ 65, and LVEF < 50% such that in each subgroup, cases and controls in had identical prevalences of the covariate. Our model was also tested in a subgroup balanced for diabetes, hypertension, and dyslipidemia with minimal decrease in model performance, suggesting the model detects CKD-specific imaging features rather than detecting cardiovascular correlates of CKD. AUC, sensitivity, and specificity for the subgroups were similar to overall model performance across all 3 clinical sites (Supplementary Table 4).*

These findings are contextualized in our **Discussion:**

> *Consistent model performance when case and control groups had identical prevalences of covariates – in isolation and in combination – suggests that the model learned shared features across high-risk groups, rather than exploiting differences in comorbidities (Supplementary Table 4).*

We also improved the aesthetics of Figure 1 at the request of Reviewer 2:



Reviewer #1: I appreciate the authors' efforts to address the concerns raised in the previous review. However, I still have several concerns.

1. The added "balanced subgroup" analyses balance only one covariate at a time, such as diabetes, hypertension, CAD, age ≥65 years, or LVEF <50%. This approach does not adequately control for the combined cardiometabolic phenotype that distinguishes CKD cases from controls. A stronger analysis would require multivariable adjustment, propensity-score matching/IPW, or comparison with a clinical baseline model including age, sex, diabetes, hypertension, CAD, LVEF, and structured echocardiographic variables.

Reply: We thank the reviewer for this astute observation. In our empirical experience, AI using the echocardiography video itself outperforms structured echocardiographic variables as the video contains pixel-level data that non-imaging structured variables cannot capture.

To better control for cardiometabolic phenotypes and augment our previous analysis, we created a multi-variate subgroup balanced for hypertension, dyslipidemia, and diabetes as well as age and sex, and show that AUC, sensitivity, and specificity are similar to overall model performance across the 3 sites. We have modified **Supplementary Table 4** to include these findings:

Site	Comorbidity	AUC	Sensitivity	Specificity
CSMC	Diabetes	0.748 (0.722, 0.773)	76.2% (72.0%, 80.5%)	58.2% (56.0%, 60.4%)
	Hypertension	0.765 (0.741, 0.790)	77.0% (72.8%, 81.0%)	61.6% (59.4%, 63.7%)
	Dyslipidemia	0.745 (0.719, 0.770)	73.0% (68.7%, 77.3%)	61.5% (59.3%, 63.6%)
	CAD	0.726 (0.698, 0.752)	72.5% (68.0%, 76.9%)	60.2% (58.0%, 62.3%)
	Age ≥ 65	0.775 (0.751, 0.799)	77.5% (73.4%, 81.6%)	63.0% (60.8%, 65.1%)
	LVEF < 50%	0.731 (0.705, 0.760)	76.0% (71.7%, 80.1%)	56.9% (54.7%, 59.0%)
	Diabetes, hypertension, dyslipidemia	0.774 (0.749, 0.797)	79.7% (75.7%, 83.5%)	59.3% (57.1%, 61.4%)
KPNC	Diabetes	0.693 (0.664, 0.720)	73.8% (69.5%, 77.9%)	50.2% (48.0%, 52.4%)
	Hypertension	0.651	60.8%	59.2%

		(0.620, 0.681)	(56.0%, 65.5%)	(57.1%, 61.4%)
	Dyslipidemia	0.704 (0.676, 0.732)	73.% (69.2%, 77.7%)	54.7% (52.6%, 57.0%)
	CAD	0.703 (0.674, 0.730)	74.8% (70.5%, 78.9%)	52.5% (50.3%, 54.7%)
	Age ≥ 65	0.742 (0.715, 0.768)	79.0% (74.9%, 82.9%)	54.9% (52.7%, 57.1%)
	LVEF < 50%	0.723 (0.693, 0.749)	80.7% (76.8%, 84.6%)	46.5% (44.3%, 48.7%)
	Diabetes, hypertension, dyslipidemia	0.700 (0.673, 0.727)	70.8% (66.3%, 75.2%)	56.5% (54.3%, 58.7%)
SHC	Diabetes	0.705 (0.666, 0.744)	65.5% (58.8%, 72.0%)	64.9% (61.9%, 67.9%)
	Hypertension	0.703 (0.664, 0.742)	65.5% (58.9%, 72.0%)	65.0% (62.0%, 67.9%)
	Dyslipidemia	0.692 (0.653, 0.729)	66.0% (59.2%, 72.4%)	64.2% (61.2%, 67.2%)
	CAD	0.701 (0.661, 0.739)	69.5% (62.9%, 75.7%)	59.9% (56.9%, 62.9%)
	Age ≥ 65	0.723 (0.685, 0.760)	67.0% (60.3%, 73.5%)	64.6% (61.7%, 67.5%)
	LVEF < 50%	0.714 (0.675, 0.752)	67.5% (61.0%, 73.9%)	62.5% (59.4%, 65.5%)
	Diabetes, hypertension, dyslipidemia	0.737 (0.684, 0.785)	68.0% (58.7%, 77.1%)	65.0% (60.8%, 69.1%)

Supplementary Table 4. *AUC, sensitivity, and specific of model performance with 95% confidence intervals in subgroups with equal prevalence of samples with and without the comorbidity of interest across all 3 clinical sites*

Our findings are described in **Results: Model Performance:**

> *We also tested our model in balanced subgroups based on diabetes, hypertension, dyslipidemia, CAD, age ≥ 65, and LVEF < 50% such that in each subgroup, cases and controls in had identical prevalences of the covariate. Our model was also tested in a subgroup balanced for diabetes, hypertension, and dyslipidemia with minimal decrease in*

model performance, suggesting the model detects CKD-specific imaging features rather than detecting cardiovascular correlates of CKD.

And contextualized in our **Discussion:**

> *Consistent model performance when case and control groups had identical prevalences of covariates – in isolation and in combination – suggests that the model learned shared features across high-risk groups, rather than exploiting differences in comorbidities (Supplementary Table 4).*

Finaly, we have modified our **STAR Methods: Statistical Analysis** to describe the subgroup construction:

> *For each covariate, balanced subgroups were defined as cases and controls with identical prevalences of each covariate; a prevalence of 50% was chosen to enrich the controls for each covariate. In the CSMC internal test cohort and KPNC external test cohort, 400 cases were randomly sampled such that for each covariate, 50% of cases were positive and 50% were negative; similarly, 2,000 controls were sampled such that for each covariate, 50% of controls were positive and 50% were negative. This process ensured that cases and controls had identical distributions of each covariate…AUC, sensitivity, and specificity with 95% confidence intervals were then calculated by bootstrapping as above. Using the methodology above, we also created a multivariate balanced subgroup using co-morbidities of diabetes mellitus, hypertension, and dyslipidemia to capture metabolic syndrome. For a combined phenotype of diabetes mellitus, hypertension, and dyslipidemia, 50% of cases were positive for all three co-morbidities and 50% were negative for all three. Given the smaller size of the SHC external test cohort and this multivariate criteria, 50 cases and 250 controls were randomly sampled.*

2. In KPNC, specificity is low in several clinically important high-risk groups, including patients with diabetes, CAD, LVEF <50%, and age ≥65 years. These findings are directly aligned with the concern that the model may generate many false positives in cardiometabolically complex populations. Therefore, the authors' statement that performance is "similar" or "consistent" across subgroups seems overly optimistic.

Reply: Thank you for this feedback. In these particular subgroups, with a decrease in specificity, there is an improvement in sensitivity with higher than average sensitivity in these subgroups. We also note that this was not apparent in the two other cohorts (CSMC nor SHC). In totality, we suspect this is due to random noise along / multiple hypothesis testing with many subgroups

assessed as recommend by the reviewer. Given the relatively lower number of these patients, it is correct that we chose an overall cut-off threshold to determine sensitivity and specificity, and more work should be done to assess whether population specific cutoffs would be needed in running this model. In this revision, we comment upon this in the **Discussion:**

> *With sensitivities from 60% to 81% our model can be leveraged for initial screening of occult or early-stage CKD (Tables 3-5), however further work is needed to optimize the decision threshold and balance sensitivity and specificity.*

3. The stage-specific results show lower performance for mild and moderate CKD and better performance for severe CKD or ESRD. This pattern is more consistent with detection of advanced systemic disease burden than with reliable early CKD detection.

Reply: We acknowledge the challenge of screening tools in that later stage disease have more severe phenotypes and thus more easily detected. That said, our analysis suggests that performance was not significantly different based on EGFR threshold (Table 2), which suggests the model would be helpful in both early as well as late CKD or ESRD. Optimal threshold for disease discrimination is important as there is a trade-off between sensitivity and specificity at various disease stages.

Table 2. Model performance on patients with laboratory evidence of CKD

eGFR threshold (mL/min/1.73 m^2)	**Cases**	**Controls**	**AUC**	**Sensitivity**	**Specificity**
30	966	4,761	0.741 (0.740, 0.743)	82.5% (82.3%, 82.7%)	48.8% (48.8%, 48.8%)
45	1,649	4,078	0.726 (0.725, 0.727)	77.1% (76.9%, 77.2%)	53.8% (53.7%, 53.9%)
60	2,503	3,224	0.733 (0.732, 0.734)	74.6% (74.5%, 74.7%)	58.4% (58.3%, 58.5%)

In this revision, we comment upon this in the **Discussion:**

> *With sensitivities from 60% to 81% our model can be leveraged for initial screening of occult or early-stage CKD (Tables 3-5), however further work is needed to optimize the decision threshold and balance sensitivity and specificity.*

4. Several numerical and reporting inconsistencies remain. For example, the severe CKD AUC is reported as 0.733 with a 95% CI of 0.755-0.790, which is mathematically impossible.

Reply: We thank the reviewer for noting this typographical error. While Supplementary Table 3 correctly lists the AUC as 0.773, we have revised our **Results: Model Performance** to:

> *Performance was consistent across stages of CKD with an AUC of 0.694 [0.670 – 0.718] in discriminating mild CKD, AUC 0.701 [0.689 – 0.712] for moderate CKD, AUC 0.773 [0.755 – 0.790] for severe CKD, and AUC 0.840 [0.829 – 0.850] for ESRD (Supplementary Table 3).*

Reviewer #2: After review, I have a few remaining minor suggestions:

1. Results-Model Performance: "AUC 0.733 [0.755 - 0.790] for severe CKD" may contain a data error. Please check this value carefully. Additionally, please re-verify all data throughout the entire text, including both the main body and the figures, to ensure accuracy.

Reply: We thank the reviewer for noting this typographical error, and we have re-verified all data throughout the text. While Supplementary Table 3 correctly lists the AUC as 0.773, our manuscript body contains a typo, and we have revised our **Results: Model Performance** as such:

Performance was consistent across stages of CKD with an AUC of 0.694 [0.670 – 0.718] in discriminating mild CKD, AUC 0.701 [0.689 – 0.712] for moderate CKD, AUC 0.773 [0.755 – 0.790] for severe CKD, and AUC 0.840 [0.829 – 0.850] for ESRD (Supplementary Table 3).

We have also reviewed the manuscript completely to address other typographical errors in the **Introduction**:

To this end, we present a deep learning model that detects CKD from parasternal long axis (PLAX) echocardiogram videos developed using data from 62,818 patients at Cedars-Sinai Medical Center (CSMC) and externally validated it in a cohort of 2,762 patients at Stanford Healthcare (SHC) and 41,611 patients from Kaiser-Permanente Northern California (KPNC).

In the **Results: Study population:**

Our model was developed using data from 62,818 patients who received care at CSMC and had 146,631 echocardiography studies with 325,377 PLAX videos.

In the **STAR Methods: Model Development and Interpretability:**

We then trained our model on one NVIDIA RTX 3090 graphics processing unit

As well as **Table 5**:

Table 5. Performance of model in discriminating any stage of CKD by patient subgroup in the SHC external validation cohort with 95% confidence intervals in parentheses.

Subgroup	Number of patients	CKD Prevalence	AUC	Sensitivity	Specificity
Diabetes mellitus	901	424 (47.1%)	0.704	68.9%	63.6%

			(0.669, 0.741)	(64.1%, 73.8%)	(59.2%, 67.9%)
Hypertension	1,392	534 (38.4%)	0.683 (0.654, 0.712)	64.8% (60.4%, 69.2%)	61.1% (57.9%, 64.4%)
Coronary artery disease	810	403 (49.8%)	0.638 (0.595, 0.680)	63.1% (57.5%, 68.8%)	54.8% (49.6%, 69.8%)
Dyslipidemia	1,484	593 (40.0%)	0.695 (0.667, 0.723)	63.3% (59.0%, 67.6%)	63.8% (60.7%, 66.9%)
EF < 50%	1,047	419 (40.0%)	0.683 (0.646, 0.720)	79.2% (75.0%, 83.3%)	45.9% (41.0%, 50.8%)
EF ≥ 50%	1,756	404 (23.0%)	0.695 (0.672, 0.718)	60.8% (57.1%, 64.6%)	67.6% (65.5%, 69.6%)
Male	1,373	497 (36.2%)	0.687 (0.656, 0.717)	60.1% (55.2%, 64.9%)	66.2% (62.9%, 69.3%)
Female	1,389	301 (21.7%)	0.695 (0.660, 0.730)	61.9% (55.8, 67.8%)	68.6% (66.0%, 71.2%)
Age < 65 years	1,352	317 (23.4%)	0.745 (0.715, 0.776)	62.4% (56.6%, 68.1%)	74.3% (71.6%, 76.9%)
Age ≥ 65 years	1,411	481 (34.1%)	0.641 (0.608, 0.673)	59.6% (54.6%, 64.7%)	60.6% (57.5%, 63.6%)

2. When discussing performance across CKD stages, the terminology sometimes varies between "mild CKD, moderate CKD, severe CKD, and ESRD" and phrases like "stage 3a CKD or worse" or "eGFR < 30/45/60 mL/min/1.73 m2". While these are related, maintaining a consistent staging terminology (e.g., always referring to CKD stages 1-5, or eGFR categories G1-G5, with clear definitions where necessary) when discussing performance across the spectrum of disease would enhance clarity and readability.

Reply: Thank you for this suggestion. Terminology has been standardized to “mild CKD, moderate CKD, severe CKD, and ESRD” as such:

> *2,182 (85.9%) patients had had measured eGFR < 90 mL/min/1.73 m^2 within 1 week of echocardiography, reflecting evidence of at least mild renal damage but not necessarily CKD. 1,202 patients had longitudinal eGFR values, with 657*

(54.7%) patients having a history of at least two measured eGFR < 90 mL/min/1.73 m^2 at least 90 days apart, indicating possible mild CKD or worse; 276 (42.0%) patients had eGFR < 60 mL/min/1.73 m^2, corresponding to mild-to-moderate CKD or worse… All 7,319 patients had signs of renal damage with eGFR < 90 mL/min/1.73 m^2, and eGFR < 60 mL/min/1.73 m^2 was measured in 1,057 (14.4%) patients…113 (6.5%) patients had at least two measured eGFR < 90 mL/min/1.73 m^2 at least 90 days apart, suggesting they had at least mild CKD… 365 (47.2%) patients had eGFR < 90 mL/min/1.73 m^2 within 1 week of echocardiography, and 79 (10.2%) patients had eGFR < 60 mL/min/1.73 m^2. Of the 773 false-positive patients, 365 patients had longitudinal creatinine values. 181 (49.6%) patients had possible mild CKD or worse with at least two measured eGFR values < 90 mL/min/1.73 m^2 at least 90 days apart, and 106 (29.0%) patients with at least mild-to-moderate CKD given at least two measured eGFR < 60 mL/min/1.73 m^2 at least 90 days apart.

3. Results: Model Performance: "The model generated predictions for 86 PLAX videos per second, demonstrating high efficiency." Since different videos may have varying frame rates and durations, the corresponding number of image frames also differs from one video to another. To avoid ambiguity arising from reporting performance as "the number of videos processed per second," please use a more objective metric—e.g., reporting the system's end-to-end throughput in frames per second (FPS), namely the number of image frames processed per second.

Reply: Thank you for this advice. Our **Results: Model Performance** has been revised to report the number of frames processed per second:

The model processed 4,015.8 frames per second, demonstrating high efficiency.

4. Figure 1 needs further aesthetic improvement. Consider using light grey blocks or similar color partitioning to enhance its appearance.

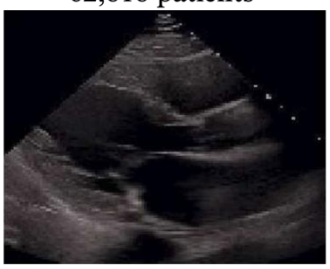
Reply: We thank the reviewer for this excellent suggestion. Figure 1 has been revised to:

Figure 1. *Schematic diagram of model development, model performance in the Cedars-Sinai internal test cohort, Kaiser-Permanente Northern California external validation cohort, and Stanford Healthcare external validation cohort and the potential utility of interpretable deep learning to screen for CKD*

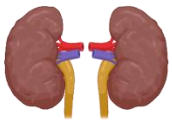
Deep Learning for CKD Detection

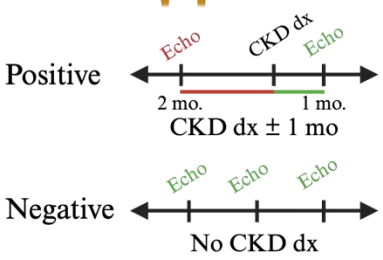
Model Input

325,377 PLAX videos from 62,818 patients

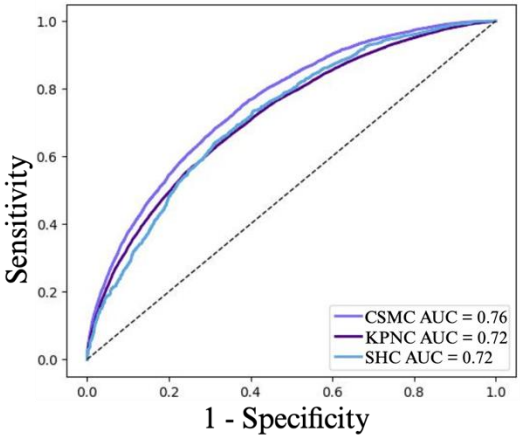


CKD Diagnosis





Model Performance

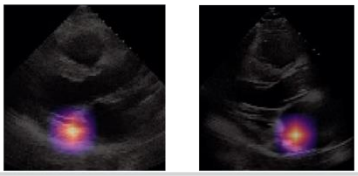


CSMC Internal Test Cohort (n=6,282)	AUC 0.76 Sensitivity 76%
KPNC External Validation Cohort (n=101,016)	AUC 0.72 Sensitivity 77%

Clinical Utility of AI for CKD Screening

PLAX

Explainable AI Model



Clinical Management

49 Spadina Ave. Suite 200
Toronto ON M5V 2J1 Canada
www.biorender.com

Confirmation of Publication and Licensing Rights - Open Access

July 25th, 2026

Subscription Type:	*Institution - Academic*
Agreement number:	*ZX2A02W21Q*
Publisher Name:	*Cell Med*
Figure Title:	*Graphical Abstract*
Citation to Use:	*Created in BioRender. Yuan, V. (2026) https://BioRender.com/9zlgg1x*

To whom this may concern,

This document ("Confirmation") hereby confirms that Science Suite Inc. dba BioRender ("BioRender") has granted the following BioRender user: Victoria Yuan ("User") a BioRender Academic Publication License in accordance with BioRender's Terms of Service and Academic License Terms ("License Terms") to permit such User to do the following on the condition that all requirements in this Confirmation are met:

1) publish their Completed Graphics created in the BioRender Services containing both User Content and BioRender Content (as both are defined in the License Terms) in publications (journals, textbooks, websites, etc.); and
2) sublicense such Completed Graphics under "open access" publication sublicensing models such as CC-BY 4.0 and more restrictive models, so long as the conditions set forth herein are fully met.

Requirements of User:

1) All Completed Graphics to be published in any publication (journals, textbooks, websites, etc.) must be accompanied by the following citation either as a caption, footnote or reference for each figure that includes a Completed Graphic:
"Created in BioRender. Yuan, V. (2026) https://BioRender.com/9zlgg1x ".
2) All terms of the License Terms including all Prohibited Uses are fully complied with. E.g. For Academic License Users, no commercial uses (beyond publication in journals, textbooks or websites) are permitted without obtaining or switching to a BioRender Industry Plan.
3) A Reader (defined below) may request that the User allow their figure to be a public template for Readers to view, copy, and modify the figure. It is up to the User to determine what level of access to grant.

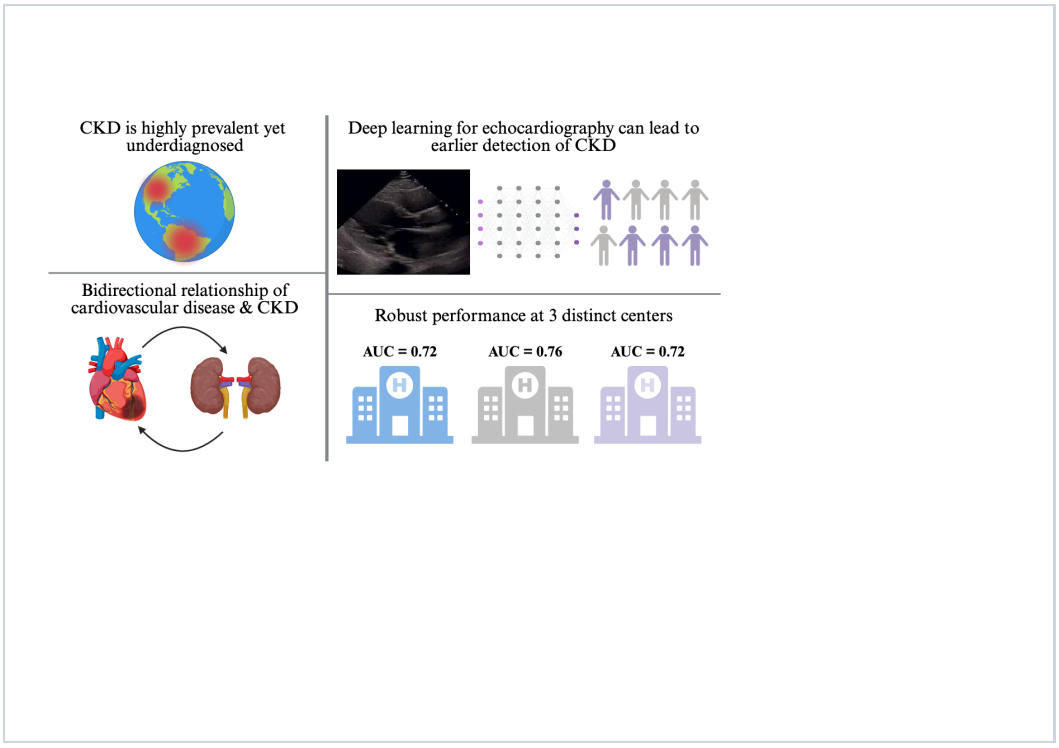
Open-Access Journal Readers:

Open-Access journal readers ("Reader") who wish to view and/or re-use a particular Completed Graphic in an Open-Access journal subject to CC-BY sublicensing may do so by clicking on the URL link in the applicable citation for the subject Completed Graphic.

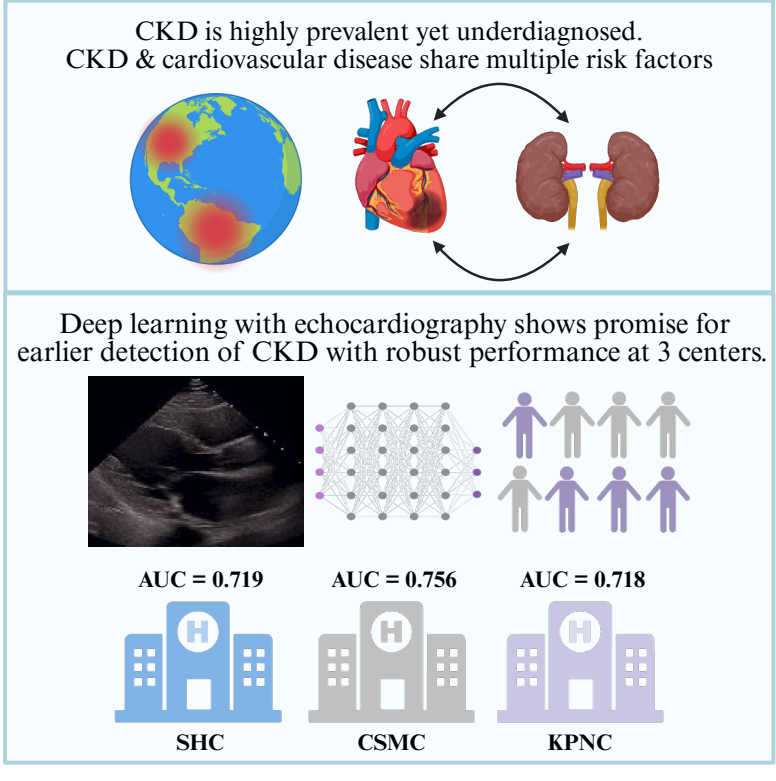
The re-use/modification options below are available after the Reader requests the User to adapt their

figure as a BioRender template and the User has granted such access.

1) View-Only/Free Plan Use: A Reader who wishes to only view the Completed Graphic may do so in the BioRender Services as either a BioRender Free Plan user or simply as a viewer. By becoming a BioRender Free Plan user, the Reader may view, modify and re-use the Completed Graphic as permitted under BioRender's Basic License Terms (e.g. personal use only, no publishing or commercial use permitted).
2) Re-Use/Publish with No Modifications: For any re-use and re-publication of a Completed Graphic with no modification(s) to the Completed Graphic made by the Reader, a Reader may do so by citing the original author using the citation noted above with the Completed Graphic. The Reader must also comply with the underlying License Terms which apply to the Completed Graphic as noted above (e.g. no commercial use for Academic License).
3) Re-Use/Publish with Modifications: For any re-use and re-publication of a Completed Graphic with a modification(s) made by the Reader, the Reader may do so by becoming a BioRender user themselves under either an Academic or Industry Plan, citing the original author using the citation noted above with the Completed Graphic and complying with the applicable License Terms.



For any questions regarding this document, or other questions about publishing with BioRender, please refer to our BioRender Publication Guide, or contact BioRender Support at support@biorender.com.



Deep Learning-Enabled Screening of Chronic Kidney Disease from Echocardiography

Victoria Yuan[1,2], Hirotaka Ieki[3], Alexander Sandhu MD[2,3], Long H. Nguyen[4], Paul P. Cheng[3], Stephanie T. Chang[3], Andrew P. Ambrosy[5], Alan C. Kwan[1], Alan S. Go MD[5], Susan Cheng MD[1], David Ouyang[1,5]

1. Department of Cardiology, Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA, United States
2. David Geffen School of Medicine, University of California, Los Angeles, CA, United States
3. Division of Cardiology, Department of Medicine, Stanford University, Palo Alto, CA, United States
4. Division of Gastroenterology, Massachusetts General Hospital, Boston, MA, United States
5. Division of Research, Kaiser Permanente Northern California, Pleasanton, CA, United States

Lead Contact and Corresponding Author:

Dr. David Ouyang

David.ouyang@kp.org

4480 Hacienda Drive

Pleasanton, CA 94588

Total word count: 3,197

Abstract

Background: Chronic kidney disease (CKD) affects nearly 850 million individuals globally. The prevalence of undiagnosed CKD is over 60%.

Methods: Taking advantage of the relationship between CKD and cardiovascular disease, we developed a deep learning (DL) model to detect CKD from parasternal long-axis (PLAX) videos using 325,377 PLAX videos from 62,818 patients at Cedars-Sinai Medical Center (CSMC). We externally validated our model in two independent cohorts of 2,762 patients at Stanford Healthcare (SHC) and 41,611 patients at Kaiser-Permanente Northern California (KPNC).

Findings: In a held-out test cohort at CSMC, our model detected any stage of CKD with an area under the curve (AUC) of 0.756 [95% confidence interval 0.749 – 0.763], with consistently strong performance in KPNC (AUC 0.718 [0.714 – 0.723]) and SHC (AUC 0.719 [0.704 – 0.735]). Our model performed well across subgroups with and without metabolic co-morbidities, suggesting that it learned imaging features specific to CKD.

Conclusions: Our DL echo model detected CKD with robust performance at two external clinical sites, offering an avenue for noninvasive screening and improved detection rates.

Funding: This work was supported by the Sarnoff Cardiovascular Research Foundation, the American Heart Association (25POST1357984, 25AHAI1487693), and the National Institutes of Health (R00HL157421, R01HL173487, R01HL173526)

Keywords: artificial intelligence; deep learning; chronic kidney disease; echocardiography; precision diagnosis

Introduction

Nearly 850 million individuals worldwide are estimated to have chronic kidney disease (CKD)[1]. Because the early stages of CKD may be asymptomatic, the prevalence of undiagnosed CKD ranges from 61.6% to 95.5%[2,3]. Low rates of detection contribute significantly to the growing burden of CKD[1,3–7]. Current screening guidelines require both serum and urine laboratory testing[8]. Although annual testing with urine albumin-to-creatinine ratio (UACR) is standard of care in patients with diabetes mellitus, screening with UACR has rates as low as 20%[9]. Additionally, such screening is rarely performed in patients with other CKD risk factors, such as hypertension or dyslipidemia[6,10]. Suboptimal detection prevents timely management and places patients at risk of disease progression and complications[11].

CKD and cardiovascular disease (CVD) have a bidirectional relationship and frequently co-occur[7,12–15],[16,17]. Patients with CKD face a higher risk of cardiac abnormalities, such as coronary artery disease, left ventricular hypertrophy, and myocardial fibrosis[7,18]. Similarly, heart disease is a risk factor for the development of CKD and contributes to its rising incidence[1,19]. Cardiovascular-kidney-metabolic (CKM) syndrome captures the interactions between CKD, CVD, and their metabolic risk factors, which can lead to excess morbidity[15]. The relationship between CKD and CVD suggests that cardiac imaging may be leveraged to detect renal dysfunction. Early detection of CKD is imperative to connect patients to cardioprotective and renoprotective therapies.

Echocardiography is a widely used, noninvasive imaging modality that is frequently performed to evaluate cardiac pathologies. Given the co-occurrence of CVD and CKD, echocardiography offers an avenue for opportunistic screening of CKD. Deep learning (DL) applied to echocardiography has demonstrated great promise to enhance image evaluation and diagnostics. Current efforts have diagnosed cardiac diseases such as valvular pathologies[20–23], amyloidosis[24,25], and cardiomyopathy[26–28], in addition to extra-cardiac pathologies such as liver disease[29] and elevated blood urea nitrogen[30]. Given its success in detecting a wide range of pathologies, DL can facilitate identification of CKD from cardiac imaging and address the need for improved detection.

To this end, we present a deep learning model that detects CKD from parasternal long axis (PLAX) echocardiogram videos developed using data from 62,818 patients at Cedars-Sinai Medical Center (CSMC) and externally validated it in a cohort of 2,762 patients at Stanford Healthcare (SHC) and 41,611 patients from Kaiser-Permanente Northern California (KPNC). The

PLAX view is commonly obtained and easily accessible in the point-of-care setting, making it an optimal view for opportunistic screening of CKD. With consistently strong performance in multiple cohorts, our model can improve CKD screening through opportunistic screening of high-risk patients and promote the use of deep learning to bridge diagnostic gaps.

Results

Study population

Our model was developed using data from 62,818 patients who received care at CSMC and had 146,631 echocardiography studies with 325,377 PLAX videos. The case cohort consisted of 10,472 patients with 24,084 studies and 51,483 PLAX videos who had an ICD-10 code for any stage of CKD and echocardiogram within 1 month of diagnosis. The control cohort included 52,346 patients who had never been diagnosed with an ICD-10 code for CKD and who had 122,547 echocardiography studies with 273,894 PLAX videos. Demographic and clinical information of the case and control cohorts are presented in Table 1, and in the training, validation, and test sets in Supplementary Table 2. The case cohort had higher prevalence of hypertension and diabetes mellitus, which are both known risk factors for CKD. LVEF, LV systolic and diastolic volumes, rate of CAD, and rate of dyslipidemia were similar between case and control cohorts. After splitting the cohort on the patient-level into train, validation, and test sets in an 8:1:1 ratio, the internal test set consisted of 6,282 patients with 14,584 studies and 32,483 PLAX videos.

The external validation cohort at Kaiser-Permanente had 41,611 patients with 46,970 echocardiography studies and 66,249 PLAX videos. Within this cohort, 10,254 patients with 11,671 echocardiography studies and 16,600 PLAX videos who had been diagnosed with any stage of CKD. 31,357 patients had never received an ICD-10 code for CKD; there were 35,299 studies with 49,649 PLAX videos across these patients. In comparison to the case cohort at CSMC, the case cohort at KPNC had higher rates of diabetes mellitus, hypertension, dyslipidemia, and CAD. The external validation cohort at SHC consisted for 798 case patients with 916 studies and 1,426 PLAX videos and 1, 964 control patients with 2,207 studies and 3,070 PLAX videos. Urine albumin values were not available for the SHC cohort, so UACR was not calculated. Case patients at SHC had a higher mean eGFR, lower rate of hypertension, and higher rates of diabetes mellitus, dyslipidemia, and CAD. Demographics were similar at all three sites.

Model Performance

Our deep learning model identified any stage of CKD with strong performance in the CSMC, KPNC, and SHC cohorts. In the CSMC internal test cohort, our algorithm detected CKD with an AUC of 0.756 [0.749 – 0.763], sensitivity of 75.6% [74.0% - 76.4%], and specificity of 61.3% [61.3% – 62.4%] (Figure 1). The model processed 4,015.8 frames per second, demonstrating high efficiency. Our model demonstrated a calibration slope of 1.15, calibration intercept of -0.02, and Brier score loss of 0.11 (Supplementary Figure 1). The model was overall well-calibrated with slight overestimation. The model distinguished CKD with strong performance across the spectrum of disease severity. Performance was consistent across stages of CKD with an AUC of 0.694 [0.670 – 0.718] in discriminating mild CKD, AUC 0.701 [0.689 – 0.712] for moderate CKD, AUC 0.773 [0.755 – 0.790] for severe CKD, and AUC 0.840 [0.829 – 0.850] for ESRD (Supplementary Table 3). When defining CKD as a persistent reduction in eGFR, the model detected patients with CKD and eGFR below 30 mL/min/1.73 m^2 with an AUC of 0.741, patients with eGFR below 45 mL/min/1.73 m^2 with an AUC of 0.726, and patients with eGFR < 60 mL/min/1.73 m^2 with an AUC of 0.733 (Table 2).

Model performance was also compared against a clinical CKD risk calculator recommended by the Center for Disease Control and Prevention to detect both any stage of CKD and to detect moderate CKD or greater (Supplementary Figure 2)[31]. Utilizing age, sex, and 7 comorbidities, the clinical risk calculator detected any stage of CKD in the CSMC cohort with an AUC of 0.64 with sensitivity 87.6% and specificity 28.3%. To detect moderate CKD or greater, the risk calculator had AUC 0.62, sensitivity 85.7%, and specificity 28.3%. In comparison, DL detected moderate CKD or higher with AUC 0.77, sensitivity 76.4%, and specificity 61.9%.

We then tested model performance in subgroups of patients with diabetes mellitus, hypertension, and hyperlipidemia, as these cohorts are considered higher risk for CKD but are not routinely screened. Our model demonstrated robust performance across subgroups with AUCs of 0.728 [0.716 – 0.739] in patients with diabetes; 0.752 [0.743 – 0.759] in patients with hypertension; 0.756 [0.745 – 0.767] in patients with CAD; and 0.752 [0.742 – 0.761] in patients with dyslipidemia (Table 3, Supplementary Figure 3). Model performance was robust to LVEF (AUC 0.700 [0.685 – 0.715] in patients with LVEF < 50% versus AUC 0.770 [0.762 – 0.779] in patients with LVEF ≥ 50%); sex (AUC 0.762 [0.750 – 0.775] in female patients versus AUC 0.752 [0.743,

0.760] in male patients); and age (AUC 0.820 [0.810 – 0.830] in patients under 65 years old versus AUC 0.708 [0.700, 0.717] in patients equal to or greater than 65 years old).

During external validation, our model detected any stage of CKD in the KPNC cohort with AUC of 0.718 [0.714 – 0.723], sensitivity 77.2% [76.6% – 77.8%], and specificity 52.5% [52.1% – 52.9%] and in the SHC cohort with AUC of 0.719 [0.704 – 0.735], sensitivity 67.1% [64.7% – 69.6%], and specificity 65.1% [63.4% – 66.8%]. Performance in detecting mild, moderate, severe, and end-stage CKD in the KPNC external validation cohort had AUCs ranging from 0.678 to 0.838 and sensitivity from 62.4% to 88.1% (Supplementary Table 3). Model performance in the KPNC and SHC cohorts was higher in subgroups with younger age < 65 (Tables 4 and 5). In the KPNC cohort, AUCs were consistent in subgroups by hypertension, CAD, and dyslipidemia and in the SHC cohort, by diabetes, hypertension, and dyslipidemia. Subgroups of patients with LVEF < 50%, age ≥ 65, and CAD had decreased performance at both SHC and KPNC. Performance also dropped in KPNC patients with diabetes (Tables 4 and 5). We also tested our model in balanced subgroups based on diabetes, hypertension, dyslipidemia, CAD, age ≥ 65, and LVEF < 50% such that in each subgroup, cases and controls in had identical prevalences of the covariate. Our model was also tested in a subgroup balanced for diabetes, hypertension, and dyslipidemia with minimal decrease in model performance, suggesting the model detects CKD-specific imaging features rather than detecting cardiovascular correlates of CKD. AUC, sensitivity, and specificity for the subgroups were similar to overall model performance across all 3 clinical sites (Supplementary Table 4).

Finally, to demonstrate the clinical utility of our model, we assessed its performance in detecting moderate to severe CKD and excluded ESRD. DL detection showed consistent performance across all 3 clinical centers with sensitivities from 70.3% – 73.8% and specificities from 52.5% – 65.1% (Supplementary Table 5).

Error Mode Analysis

We analyzed patients whom the AI model identified patients as having CKD, but who did not have a corresponding ICD-10 code. In the CSMC internal test cohort, 2,778 (44.2%) of patients were identified as having CKD based on their PLAX video without a clinical diagnosis. Of these patients, 2,541 patients had eGFR measured within 1 week of echocardiography, and 2,182 (85.9%) patients had had measured eGFR < 90 mL/min/1.73 m^2 within 1 week of echocardiography, reflecting evidence of at least mild renal damage but not necessarily CKD. 1,202 patients had

longitudinal eGFR values, with 657 (54.7%) patients having a history of at least two measured eGFR < 90 mL/min/1.73 m^2 at least 90 days apart, indicating possible mild CKD or worse; 276 (42.0%) patients had eGFR < 60 mL/min/1.73 m^2, corresponding to mild-to-moderate CKD or worse. Using their eGFR and UACR measurements, we found that 111 (4.0%) patients had KDIGO risk scores corresponding with moderately to very high increased risk.

Similarly, in the KPNC external validation cohort, 16,019 (38.5%) patients were identified by the model as having CKD but did not have a clinical diagnosis. 7,319 (45.7%) patients had eGFR measured within 1 week of their study. All 7,319 patients had signs of renal damage with eGFR < 90 mL/min/1.73 m^2, and eGFR < 60 mL/min/1.73 m^2 was measured in 1,057 (14.4%) patients. 1,737 patients had longitudinal creatinine values measured, with 113 (6.5%) patients had at least two measured eGFR < 90 mL/min/1.73 m^2 at least 90 days apart, suggesting they had at least mild CKD. Additionally, 426 (2.7%) patients had albuminuria with UACR ≥ 30 within one month of their PLAX video. KDIGO risk scores could be calculated for 557 (3.5%) patients, with 251 (45.1%) patients having moderate to very high risk KDIGO scores. In the SHC cohort, 773 patients were identified by the model as having CKD although they did not have a clinical diagnosis. 365 (47.2%) patients had eGFR < 90 mL/min/1.73 m^2 within 1 week of echocardiography, and 79 (10.2%) patients had eGFR < 60 mL/min/1.73 m^2. Of the 773 false-positive patients, 365 patients had longitudinal creatinine values. 181 (49.6%) patients had possible mild CKD or worse with at least two measured eGFR values < 90 mL/min/1.73 m^2 at least 90 days apart, and 106 (29.0%) patients with at least mild-to-moderate CKD given at least two measured eGFR < 60 mL/min/1.73 m^2 at least 90 days apart. UACR values were not available for patients at SHC, and KDIGO risk scores were not calculated.

We then identified patients whom the AI model identified as not having CKD, but who did have a corresponding ICD-10 code. In the CSMC internal test cohort, 488 (7.8%) patients with 964 PLAX videos had a clinical diagnosis of CKD but were not detected by the model. Of these videos, 515 (53.4%) were associated with a diagnosis of moderate CKD while 115 (11.9%) videos were associated with mild CKD, 149 (15.5%) with severe CKD, and 185 (19.2%) with ESRD. Similarly, in the Stanford external test cohort, 289 (10.5%) patients with 778 PLAX videos in the case cohort did not have a model prediction of CKD. 390 (50.1%) PLAX videos were from patients with a diagnosis of moderate CKD, and 274 (35.2%) videos from those with mild CKD. In the KPNC cohort, 2,939 (7.1%) patients with 3,796 videos did not have CKD by DL, despite having

a clinical diagnosis. Of these 2,939 patients, 2,570 (87.4%) patients with 3,293 were staged. 2,464 (74.8%) videos classified as false negative predictions were associated with moderate CKD, while 295 (9.0%) videos were associated with mild CKD, 258 (7.8%) with severe CKD, and 276 (8.4%) with ESRD.

Model Interpretability

To support explainability for clinical use, we generated GradCAM heatmaps of significant pixels for CKD detection. GradCAM highlighted the mitral valve and left atrium as important features for AI model decision-making (Figure 2). These results are consistent with existing literature highlighting excess prevalence of mitral valve dysfunction that occurs in the early stages of CKD, with patients with CKD having higher rates of mitral regurgitation and undergoing mitral valve calcification 10-20 years earlier than those without CKD[32–34].

Discussion

In this work, we developed a deep learning model to detect chronic kidney disease (CKD) from parasternal long axis (PLAX) echocardiogram videos. Our model demonstrated strong performance to identify any stage of CKD across diverse patient subgroups. We validated our model in two geographically distinct centers, emphasizing its ability to generalize across different patient populations and clinical protocols. Moreover, our model detected patients with CKD when labelled by laboratory criteria instead of ICD-10 codes, suggesting that it is robust to label shift and label noise. With an overall sensitivity of 76%, our study offers an innovative avenue to opportunistically screen for CKD in patients undergoing echocardiography. Improved detection can connect patients to treatment earlier in their disease course and decrease the growing burden of CKD.

CKD affects 850 million individuals globally and is projected to be the 5th leading cause of lost years of life by 2040[1]. The rising prevalence of CKD reflects a rise in its metabolic risk factors of hypertension, diabetes, and obesity[1,15]. These factors also increase risk of CVD, with renal dysfunction contributing additive risk[7,14,15]. This interplay between metabolic, renal, and cardiovascular disease is captured by cardiovascular-kidney-metabolic (CKM) syndrome, which in turn increases morbidity and premature mortality. Early recognition of CKD and CKM can connect patients to appropriate management and slow disease progression. Our analysis of

CSMC's historical echocardiography cohort revealed that 57.3% of patients did not have sufficient eGFR measurements to diagnose CKD. Yet many of these patients are at higher risk of CKD due to cardiovascular disease. There is a critical need to improve identification of CKD syndrome. However, current methods to detect CKD have low screening rates.

Existing DL studies predict risk of CKD using demographic, clinical, and laboratory values[35–39]. While these models show promise for detection, they require serum values or physical exam findings. When compared to a recommended clinical risk calculator, DL-based detection can improve specificity and detect earlier stages of CKD. Other models predict progression to ESRD in patients with established CKD, rather than detecting new-onset CKD[38,39]. Holmstrom, et al. utilized 12-lead electrocardiograms to detect CKD within a 1-year window[40]. Identifying CKD within a narrower window can improve disease detection and facilitate timely treatment to prevent CKD progression.

Our model offers a noninvasive, rapid avenue to detect CKD from PLAX videos with high fidelity. Performance was consistent in high-risk patients with hypertension and dyslipidemia across all three sites (Tables 3-5) and in patients with diabetes at two sites (Tables 3-5). Consistent model performance when case and control groups had identical prevalences of covariates – in isolation and in combination – suggests that the model learned shared features across high-risk groups, rather than exploiting differences in comorbidities (Supplementary Table 4). Deep learning detection also demonstrated stronger performance in patients under 65, suggesting that it can serve as a screening tool for younger patients who face higher mortality from CKD[41]. Because our DL-based screening analyzes echocardiography, it may be potentially easier to use in a clinical setting, whereas the necessary variables for other calculators may not always be obtained from health records. Moreover, our work detects any stage of CKD to facilitate earlier detection, which can complement existing studies that focus on progression to ERSD[38,39,42].With sensitivities from 60% to 81% our model can be leveraged for initial screening of occult or early-stage CKD (Tables 3-5); however further work is needed to optimize the decision threshold and balance sensitivity and specificity. Combined with clinical histories, our DL algorithm may address diagnostic gaps for high-risk patients, for whom screening is not regularly performed[6,9]. Our model also identified patients without a clinical diagnosis of CKD, but based on their eGFR, UACR, and KDIGO risk score, may have subclinical or early-stage CKD and may benefit from intervention. With the American Heart Association's recent recommendations to screen for and stage CKM syndrome,

we envision that our deep learning model can be integrated into clinical workflows for simultaneous evaluation of cardiac and renal function[15]. Early identification of CKD can enable earlier intervention, improve patient awareness, and slow disease progression.

Explainability analysis with GradCAM highlights that our algorithm focused on the mitral valve and left atrium, two structures whose dysfunction may be early manifestations of CKD-induced cardiac disease[34,43]. Our model identified imaging biomarkers that can be investigated further as part of the pathophysiology of CKD[34,43].

Future work involves validation of our model at additional sites and with different ultrasound vendors, as well as integration of our model into clinical workflows to characterize its clinical utility. Future work also includes development of a multiclassification model with a balanced training cohort. With the advent of multi-view echocardiography models, future directions also include developing a model that integrates multiple views to detect CKD and comparing its performance to a single-view model. Future research directions also include predicting CKD progression using PLAX videos, which can improve risk stratification. Given the need for CKD screening in high-risk patient populations, we anticipate that DL-based detection can increase detection rates, facilitate early intervention for patients, and help decrease the global disease burden of CKD.

In conclusion, we developed and externally validated a deep learning model to detect CKD from PLAX echocardiography videos. Our algorithm can detect any stage of CKD, including early stages that may go undiagnosed, suggesting deep learning screening for CKD from echocardiography can help bolster low detection rates and promote more timely management for physicians.

Limitations of Study

Limitations of the study include its retrospective analysis of data from tertiary academic centers, which may introduce bias in patient selection. Additionally, labels for CKD were extracted using ICD-10 codes, which is intended for clinical billing and may be imprecise. The study cohort had a decreased prevalence of mild and moderate CKD, which may introduce biases in disease recognition. While the model had a modest false negative rate of 7.1-10.5%, a higher proportion of videos were associated with diagnoses of moderate CKD. While KDIGO risk assessments could be calculated for the CSMC and KPNC cohorts, data on UACR was not available in the SHC cohort, which may limit the conclusions from this analysis.

Resource Availability

Lead Contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Dr. David Ouyang (David.ouyang@kp.org)

Materials Availability: This study did not generate new unique reagents.

Data and Code Availability

Model weights and code are publicly available on Github at https://github.com/echonet/ckd. Due to its potentially identifiable nature, the imaging dataset will not be released.

Acknowledgements

Ms. Yuan gratefully acknowledges funding from the Sarnoff Cardiovascular Research Foundation. Dr. Ieki gratefully acknowledges funding from the American Heart Association (25POST1357984). Dr. Ouyang gratefully acknowledges support from the National Institutes of Health (R00HL157421, R01HL173487 and R01HL173526) and the American Heart Association (25AHA I1487693).

Author Contributions

V.Y. and D.O. were responsible for conceptualization, methodology, and investigation V.Y. wrote the main manuscript text and prepared figures. H.I., P.P.C, A.P.A., and A.S.G. assisted with external validation. V.Y., H.I., A.S, L.H.N, P.P.C, S.T.C, A.P.A., A.C.K., A.S.G., S.C., and D.O. reviewed the manuscript. D.O. supervised the project. All authors read and approved the final article and take responsibility for its content.

Declaration of Interests: Dr. Chang reports research support from Siemens Healthcare. Dr. Ambrosy reports research support through grants from the National Institutes of Health (R01HL173866 and R01AG091005), the Food and Drug Administration (U01FD008704), The American Heart Association, The Permanente Medical Group, Kaiser Permanente Northern California Community Benefits Programs, Garfield Memorial Fund, Abbott Laboratories, Abiomed, Amarin Pharma, Inc., Bayer Pharma AG, Boehringer Ingelheim, Cordio Medical, Edwards Lifesciences LLC, Esperion Therapeutics, Inc., Merck, and Novartis. Consulting for Bristol Meyer Squibb, Bayer Pharma AG, Corstasis, Merck, Novo Nordisk, Pharmacosmos A/S,

Pfizer, scPharma, SQ Innovation, Reprieve Cardiovascular, and VisCardia. Dr. Go has received research support from the National Heart, Lung and Blood Institute and the National Institute of Diabetes, Digestive and Kidney Diseases. Dr. Ouyang reports consulting, equity, or honoraria for lectures from AstraZeneca Alexion, EchoIQ, Ultromics, Pfizer, InVision, the Korean Society of Echocardiography, and the Japanese Society of Echocardiography. The remaining authors have nothing to declare.

Figure 1. Schematic diagram of model development, model performance in the Cedars-Sinai internal test cohort, Kaiser-Permanente Northern California external validation cohort, and Stanford Healthcare external validation cohort and the potential utility of interpretable deep learning to screen for CKD. Created in BioRender, https://BioRender.com/m6w0sz2

Figure 2. Grad-CAM[44] activation was primarily at the mitral valve in PLAX videos

Table 1. Demographics and co-morbidities of the internal development cohort at Cedars-Sinai Medical Center (CSMC), the external validation cohort at Kaiser-Permanente Northern California (KPNC), and the external validation cohort at Stanford Healthcare (SHC)

	CSMC Controls (n=52,346)	**CSMC Cases (n=10,472)**	**KPNC Controls (n=31,357)**	**KPNC Cases (n=10,254)**	**SHC Controls (n=1,964)**	**SHC Cases (n=798)**
Age	67.7 ± 15.6	66.5 ± 15.2	69.2 ± 12.9	73.1 ± 13.1	61.0 ± 17.4	66.3 ± 14.9
Sex						
Male	33,645 (64.3%)	6,731 (64.3%)	15,763 (50.3%)	5,621 (54.8%)	876 (44.6%)	497 (62.3%)
Female	18,701 (35.7%)	3,741 (35.7%)	15,572 (49.7%)	4,631 (45.2%)	1088 (55.4%)	301(37.7%)
Race and Ethnicity						
American Indian or Alaska Native	106 (0.2%)	28 (0.3%)	154 (0.5%)	62 (0.6%)	7 (0.4%)	2 (0.3%)
Asian	3,594 (7.0%)	895 (8.6%)	4,695 (15.0%)	1,805 (17.6%)	384 (19.6%)	119 (14.9%)
Black or African American	6,072 (11.8%)	2,238 (21.4%)	2,191 (7.0%)	1,504 (14.7%)	103 (5.2%)	54 (6.8%)
Native Hawaiian/Pacific Islander	142 (0.3%)	46 (0.4%)	219 (0.7%)	167 (1.6%)	32 (1.6%)	23 (2.9%)
Hispanic	5,738 (11.2%)	1,715 (16.4%)	2,923 (9.3%)	1,154 (11.3%)	198 (10.1%)	111 (13.9%)
Non-Hispanic	44,390 (86.3%)	8,652 (82.6%)	28,434 (90.7%)	9,100 (88.7%)	1,635 (83.2%)	672 (84.2%)
White	36,612 (71.1%)	6,484 (61.9%)	17,152 (54.7%)	5,454 (53.2%)	1,021 (52.0%)	437 (54.8%)
Other	3,702 (7.2%)	664 (6.3%)	185 (0.5%)	37 (0.3%)	309 (15.7%)	143 (17.9%)
Unknown	2,118 (4.0%)	98 (0.9%)	191 (0.6%)	14 (0.1%)	108 (5.5%)	20 (2.5%)

Left ventricular ejection fraction %	52.8% ± 17.4%	56.9% ± 14.9%	57.9% ± 10.2%	53.5% ± 13.4%	57.9% ± 11.0%	51.0% ± 15.4%
Left ventricular end-diastolic volume index (mL/m^2)	46.6 ± 25.8	51.9 ± 27.8	49.7 ± 17.1	55.7 ± 23.8	45.8 ± 18.1	55.7 ± 30.4
Left ventricular end-systolic volume index (mL/m^2)	21.5 ± 17.4	25.9 ± 19.7	21.7 ± 12.7	27.7 ± 20.2	20.3 ± 14.1	30.3 ± 27.5
Estimated glomerular filtration rate (eGFR, $mL/min/1.73m^2$)	74.02 ± 26.6	33.3 ± 22.5	58.5 ± 5.7	38.1 ± 18.2	89.4 ± 21.0	57.1 ± 28.5
Urine albumin-to-creatinine ratio (uACR)	154.3 ± 492.2	379.3 ± 737.1	46.0 ± 86.4	129.2 ± 131.7	N/A	N/A
Coronary artery disease	12,317 (23.5%)	4,013 (38.3%)	13,497 (43.0%)	6,902 (67.3%)	407 (20.7%)	403 (50.5%)
Hypertension	22,572 (43.1%)	8,334 (79.6%)	23,436 (74.7%)	10,009 (97.6%)	858 (43.7%)	534 (66.9%)
Diabetes mellitus	9,058 (17.3%)	5,089 (48.6%)	9,579 (30.5%)	7,092 (69.1%)	477 (24.3%)	424 (53.1%)
Dyslipidemia	16,739 (32.0%)	5,171 (49.4%)	20,857 (66.5%)	9,021 (88.0%)	891 (45.4%)	593 (74.3%)

Table 2. Model performance on patients with laboratory evidence of CKD

eGFR threshold (mL/min/1.73 m^2)	**Cases**	**Controls**	**AUC**	**Sensitivity**	**Specificity**
30	966	4,761	0.741 (0.740, 0.743)	82.5% (82.3%, 82.7%)	48.8% (48.8%, 48.8%)
45	1,649	4,078	0.726 (0.725, 0.727)	77.1% (76.9%, 77.2%)	53.8% (53.7%, 53.9%)
60	2,503	3,224	0.733 (0.732, 0.734)	74.6% (74.5%, 74.7%)	58.4% (58.3%, 58.5%)

Table 3. Performance of model in discriminating any stage of CKD by patient subgroup in the CSMC internal test cohort with 95% confidence intervals in parentheses.

Subgroup	Number of patients	CKD Prevalence	AUC	Sensitivity	Specificity
Diabetes mellitus	1,438	508 (35.3%)	0.728 (0.716, 0.739)	77.4% (76.2%, 79.3%)	52.8% (50.8%, 53.4%)
Hypertension	3,086	835 (27.1%)	0.752 (0.743, 0.759)	75.0% (74.2%, 76.8%)	60.9% (59.4%, 61.1%)
Coronary artery disease	1,639	395 (24.1%)	0.756 (0.745, 0.767)	76.2% (74.9%, 78.4%)	59.8% (58.1%, 60.2%)
Dyslipidemia	2,195	500 (22.8%)	0.752 (0.742, 0.761)	72.8% (71.7%, 75.0%)	63.1% (61.7%, 63.5%)
EF < 50%	1,278	444 (34.7%)	0.700 (0.685, 0.715)	79.9% (77.9%, 81.9%)	45.6% (44.2%, 47.0%)
EF ≥ 50%	4,976	717 (14.4%)	0.770 (0.762, 0.779)	72.9% (71.3%, 74.5%)	66.9% (66.3%, 67.6%)
Male	4,043	652 (16.1%)	0.752 (0.743, 0.760)	75.4% (73.9%, 76.8%)	60.5% (59.7%, 61.2%)
Female	2,239	361 (16.1%)	0.762 (0.750, 0.775)	74.8% (72.7%, 76.9%)	64.0% (63.1%, 64.9%)
Age < 65 years	2,450	361 (14.7%)	0.820 (0.810, 0.830)	79.9% (78.2%, 81.6%)	68.0% (67.1%, 68.9%)
Age ≥ 65 years	3,948	652 (16.5%)	0.708 (0.700, 0.717)	72.1% (70.4%, 73.7%)	57.7% (56.9%, 58.4%)

Table 4. Performance of model in discriminating any stage of CKD by patient subgroup in the KPNC external validation cohort with 95% confidence intervals in parentheses.

Subgroup	Number of patients	CKD Prevalence	AUC	Sensitivity	Specificity
Diabetes mellitus	16,671	7,092 (42.5%)	0.684 (0.678, 0.691)	81.1% (80.4%, 81.8%)	40.8% (40.0%, 41.5%)
Hypertension	33,445	10,009 (29.9%)	0.691 (0.687, 0.696)	78.1% (77.4%, 78.7%)	46.4% (45.9%, 46.9%)
Coronary artery disease	20,399	6,902 (33.8%)	0.698 (0.693, 0.704)	80.9% (80.2%, 81.6%)	43.6% (42.9%, 44.2%)
Dyslipidemia	29,878	9,021 (30.2%)	0.706 (0.701, 0.711)	78.6% (77.9%, 79.2%)	48.1% (47.6%, 48.7%)
EF < 50%	1,915	711 (37.1%)	0.659 (0.635, 0.675)	86.4% (83.6%, 87.7%)	30.5% (29.5%, 33.8%)
EF ≥ 50%	8,431	1,682 (20.0%)	0.716 (0.706, 0.729)	74.4% (72.5%, 76.0%)	54.8% (54.1%, 56.1%)
Male	21,384	5,625 (26.3%)	0.713 (0.707, 0.719)	78.6% (77.8%, 79.4%)	50.0% (49.3%, 50.6%)
Female	20,203	4,634 (23.0%)	0.725 (0.719, 0.732)	75.8% (74.8%, 76.8%)	55.1% (55.5%, 55.7%)
Age < 65 years	11,947	2,207 (18.5%)	0.786 (0.777, 0.794)	77.8% (76.5%, 79.2%)	62.6% (61.8%, 63.4%)
Age ≥ 65 years	29,742	8,078 (27.2%)	0.690 (0.684, 0.695)	77.2% (76.5%, 77.9%)	47.9% (47.4%, 48.5%)

Table 5. Performance of model in discriminating any stage of CKD by patient subgroup in the SHC external validation cohort with 95% confidence intervals in parentheses.

Subgroup	Number of patients	CKD Prevalence	AUC	Sensitivity	Specificity
Diabetes mellitus	901	424 (47.1%)	0.704 (0.669, 0.741)	68.9% (64.1%, 73.8%)	63.6% (59.2%, 67.9%)
Hypertension	1,392	534 (38.4%)	0.683 (0.654, 0.712)	64.8% (60.4%, 69.2%)	61.1% (57.9%, 64.4%)
Coronary artery disease	810	403 (49.8%)	0.638 (0.595, 0.680)	63.1% (57.5%, 68.8%)	54.8% (49.6%, 69.8%)
Dyslipidemia	1,484	593 (40.0%)	0.695 (0.667, 0.723)	63.3% (59.0%, 67.6%)	63.8% (60.7%, 66.9%)
EF < 50%	1,047	419 (40.0%)	0.683 (0.646, 0.720)	79.2% (75.0%, 83.3%)	45.9% (41.0%, 50.8%)
EF ≥ 50%	1,756	404 (23.0%)	0.695 (0.672, 0.718)	60.8% (57.1%, 64.6%)	67.6% (65.5%, 69.6%)
Male	1,373	497 (36.2%)	0.687 (0.656, 0.717)	60.1% (55.2%, 64.9%)	66.2% (62.9%, 69.3%)
Female	1,389	301 (21.7%)	0.695 (0.660, 0.730)	61.9% (55.8, 67.8%)	68.6% (66.0%, 71.2%)
Age < 65 years	1,352	317 (23.4%)	0.745 (0.715, 0.776)	62.4% (56.6%, 68.1%)	74.3% (71.6%, 76.9%)
Age ≥ 65 years	1,411	481 (34.1%)	0.641 (0.608, 0.673)	59.6% (54.6%, 64.7%)	60.6% (57.5%, 63.6%)

STAR Methods

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Model weights	This paper	https://github.com/echonet/ckd
Software and algorithms		
Python v3.10.12	Python	https://www.python.org/downloads/release/python-31012/
R(2+1)D convolutional neural network	Tran, et al.	https://docs.pytorch.org/vision/main/models/generated/torchvision.models.video.r2plus1d_18.html
Gradient-weighted Class Activation Mapping	Selvaraju, et al.	https://github.com/ramprs/grad-cam/

Experimental Models and Study Participant Details

This study was approved by the Institutional Review Boards at Cedars-Sinai Medical Center, Stanford University, and Kaiser-Permanente Northern California. Participant information on age, sex, race/ancestry, and ethnicity were self-reported. Information on gender and socioeconomic status was not collected.

Data Sources and Study Population

To develop our deep learning model, we retrospectively analyzed a cohort of patients who received care at Cedars-Sinai Medical Center (CSMC) between 2016 and 2022. Similarly, we identified two patient cohorts for external validation – one at Kaiser-Permanente Northern California (KPNC) with patients who received care between 2022 and 2025 and a second at Stanford Healthcare (SHC) with patients who received care between 2010 and 2018. In both the development and external validation cohorts, the case cohort consisted of patients who were diagnosed with CKD by International Classification of Disease (ICD)-10 code (Supplementary Table 1) and who had a PLAX echocardiogram within 1 month of diagnosis. Control patients were matched to cases by age within 10 years and sex in a 1:5 ratio. For patients with a creatinine measurement within one week of echocardiogram acquisition, estimated glomerular filtration rate

(eGFR) was calculated using the 2021 CKD-EPI eGFRcr equation[45]. Urine albumin and creatinine labs within one month of echocardiogram acquisition were extracted to calculate UACR. Demographic and clinical co-morbidities were extracted from the electronic health records using ICD-10 codes. Echocardiographic parameters of left ventricular ejection fraction (LVEF) and LV volumes were obtained from the final study reports.

Methods Details

Model Development and Interpretability

We trained an R(2+1)D[46] convolutional neural network as a binary classifier to predict the presence of CKD within 1 month of PLAX echocardiogram acquisition using Python v3.10.12. R(2+1)D is an established convolutional neural network that leverages spatiotemporal convolutional blocks for video analysis. Model weights were randomly initialized at the start of training. The dataset was split into train, validation, and test sets in an 8:1:1 ratio on the patient-level. If a patient had multiple PLAX videos, each video was treated as an independent sample. Data augmentation was performed by applying an affine transformation to each video and by selecting a random start time. We then trained our model on one NVIDIA RTX 3090 graphics processing unit with a learning rate of 1e-3, batch size of 32, Adam optimization with an epsilon of 1e-8 and weight decay of 1e-2, and binary cross-entropy loss. Model training was conducted for a maximum of 100 epochs. Early stopping was utilized if the validation loss was stable for 15 consecutive epochs. The final model was selected based on the epoch with the minimum validation loss. Model interpretability was analyzed with Gradient-weighted Class Activation Mapping (Grad-CAM)[44] to generate heatmaps of pixel importance for predictions.

Quantification and Statistical Analysis

Model performance was assessed using area under the receiver operator characteristic curve (AUC), sensitivity, and specificity on the test set, which was not seen during model development and training. The 95% confidence interval (CI) was calculated by bootstrapping with 10,000 samples. Model calibration was also analyzed by quantifying the calibration slope and intercept, and Brier score loss was calculated. We evaluated performance to detect mild, moderate, severe, and end-stage CKD as defined by ICD-10 codes (Supplementary Table 1) using AUC, sensitivity, and specificity as metrics. Patients who did not have a documented stage of CKD were excluded from this analysis. The model's ability to discriminate patients with laboratory evidence

of CKD was also evaluated. We defined case patients as those with at least 2 eGFR measurements that were below a pre-defined threshold, that were taken at least 90 days apart, and that were obtained at least 6 months prior to the patient's first PLAX video. We tested thresholds of eGFR < 30, 45, and 60 mL/min/1.73 m^2. Control patients were those who did not have a history of reduced eGFR prior to echocardiography. AUC, sensitivity, and specificity with 95% confidence intervals were calculated for each threshold.

Subgroup analysis was performed based on co-morbidities of diabetes mellitus, hypertension, dyslipidemia, and coronary artery disease; sex; age greater than or equal to 65 years old versus lower than 65 years old at the time of echocardiogram acquisition; and cardiac function based on LVEF greater than or equal to 50% versus less than 50%. Balanced subgroups based on age, LVEF, and each comorbidity were then created to characterize the effect of covariates on model performance. For each covariate, balanced subgroups were defined as cases and controls with identical prevalences of each covariate; a prevalence of 50% was chosen to enrich the controls for each covariate. In the CSMC internal test cohort and KPNC external test cohort, 400 cases were randomly sampled such that for each covariate, 50% of cases were positive and 50% were negative; similarly, 2,000 controls were sampled such that for each covariate, 50% of controls were positive and 50% were negative. This process ensured that cases and controls had identical distributions of each covariate. Given the smaller size of the SHC external test cohort, 200 cases and 1,000 controls were randomly sampled with identical prevalences of each covariate. AUC, sensitivity, and specificity with 95% confidence intervals were then calculated by bootstrapping as above. Using the methodology above, we also created a multivariate balanced subgroup using co-morbidities of diabetes mellitus, hypertension, and dyslipidemia to capture metabolic syndrome. For a combined phenotype of diabetes mellitus, hypertension, and dyslipidemia, 50% of cases were positive for all three co-morbidities and 50% were negative for all three. Given the smaller size of the SHC external test cohort and this multivariate criteria, 50 cases and 250 controls were randomly sampled.

Error Mode Analysis

We identified patients who were classified as having CKD by the model but did not have a corresponding ICD-10 code, and patients who had an ICD-10 code for CKD but were not identified by the model. We analyzed their UACR performed within 31 days of echocardiography.

eGFR measurements for patients over 3 months to 1 year were obtained and compared against diagnostic criteria for CKD. Finally, the Kidney Disease: Improving Global Outcomes (KDIGO) risk score was also calculated for patients who had a recorded creatinine and UACR measurement. Patients who had a clinical diagnosis of CKD but were not identified by the model as having the disease were identified, and their disease severity was analyzed across all three clinical sites.

References

1. Francis, A., Harhay, M.N., Ong, A.C.M., Tummalapalli, S.L., Ortiz, A., Fogo, A.B., Fliser, D., Roy-Chaudhury, P., Fontana, M., Nangaku, M., et al. (2024). Chronic kidney disease and the global public health agenda: an international consensus. Nat. Rev. Nephrol. *20*, 473–485.
2. Tangri, N., Moriyama, T., Schneider, M.P., Virgitti, J.B., De Nicola, L., Arnold, M., Barone, S., Peach, E., Wittbrodt, E., Chen, H., et al. (2023). Prevalence of undiagnosed stage 3 chronic kidney disease in France, Germany, Italy, Japan and the USA: results from the multinational observational REVEAL-CKD study. BMJ Open *13*, e067386.
3. Sundström, J., Bodegard, J., Bollmann, A., Vervloet, M.G., Mark, P.B., Karasik, A., Taveira-Gomes, T., Botana, M., Birkeland, K.I., Thuresson, M., et al. (2022). Prevalence, outcomes, and cost of chronic kidney disease in a contemporary population of 2·4 million patients from 11 countries: The CaReMe CKD study. Lancet Reg. Health Eur. *20*, 100438.
4. Farrell, D.R., and Vassalotti, J.A. (2024). Screening, identifying, and treating chronic kidney disease: why, who, when, how, and what? BMC Nephrol. *25*, 34.
5. Cordero, L., and Ortiz, A. (2025). Albuminuria-based universal screening for CKD should be implemented now in high-income countries. Kidney Int. https://doi.org/10.1016/j.kint.2025.04.037.
6. Alfego, D., Ennis, J., Gillespie, B., Lewis, M.J., Montgomery, E., Ferrè, S., Vassalotti, J.A., and Letovsky, S. (2021). Chronic kidney disease testing among at-risk adults in the U.s. remains low: Real-world evidence from a national laboratory database. Diabetes Care *44*, 2025–2032.
7. Jankowski, J., Floege, J., Fliser, D., Böhm, M., and Marx, N. (2021). Cardiovascular disease in chronic kidney disease: Pathophysiological insights and therapeutic options. Circulation *143*, 1157–1172.
8. Lamprea-Montealegre, J.A., Shlipak, M.G., and Estrella, M.M. (2021). Chronic kidney disease detection, staging and treatment in cardiovascular disease prevention. Heart *107*, 1282–1288.
9. Tuttle, K.R., Wong, L., St Peter, W., Roberts, G., Rangaswami, J., Mottl, A., Kliger, A.S., Harris, R.C., Gee, P.O., Fowler, K., et al. (2022). Moving from evidence to implementation of breakthrough therapies for diabetic kidney disease. Clin. J. Am. Soc. Nephrol. *17*, 1092–1103.
10. Tuttle, K.R., and Lam, C.S.P. (2023). Kidney-heart connection: Implications of chronic kidney disease for screening and prevention of heart failure. JACC Heart Fail. *11*, 538–540.
11. Gong, J., Vaduganathan, M., and Wadhera, R.K. (2025). Chronic kidney disease prevalence and awareness among US adults. JAMA Cardiol. https://doi.org/10.1001/jamacardio.2025.4581.
12. Colombijn, J.M.T., Idema, D.L., van Beem, S., Blokland, A.M., van der Braak, K., Handoko, M.L., In 't Veld, L.F.H., Kaul, T., Kolagasigil-Akdemir, N., Kusters, M.P.T., et al. (2024). Representation of patients with chronic kidney disease in clinical trials of cardiovascular disease medications: A systematic review. JAMA Netw. Open *7*, e240427.
13. Tangri, N., Kushner, P.R., Barone, S., Arnold, M., Chen, H. (tony), and Alvarez, C.S. (2023). Undiagnosed early CKD in patients with hypertension and cardiovascular disease in the United States. J. Am. Soc. Nephrol. *34*, 372–372.
14. Burnier, M., and Damianaki, A. (2023). Hypertension as cardiovascular risk factor in chronic kidney disease. Circ. Res. *132*, 1050–1063.

15. Ndumele, C.E., Rangaswami, J., Chow, S.L., Neeland, I.J., Tuttle, K.R., Khan, S.S., Coresh, J., Mathew, R.O., Baker-Smith, C.M., Carnethon, M.R., et al. (2023). Cardiovascular-kidney-metabolic health: A presidential advisory from the American heart association. Circulation *148*, 1606–1635.
16. Hayer, M.K., Radhakrishnan, A., Price, A.M., Liu, B., Baig, S., Weston, C.J., Biasiolli, L., Ferro, C.J., Townend, J.N., Steeds, R.P., et al. (2020). Defining myocardial abnormalities across the stages of Chronic Kidney Disease: A cardiac magnetic resonance imaging study. JACC Cardiovasc. Imaging *13*, 2357–2367.
17. Christensen, J., Landler, N.E., Olsen, F.J., Sørensen, I.M.H., Bjergfelt, S.S., Ballegaard, E.L.F., Feldt-Rasmussen, B., Hansen, D., Kamper, A.-L., Christoffersen, C., et al. (2024). The role of aetiology in cardiac manifestations of chronic kidney disease: the CPH-CKD ECHO study. Int. J. Cardiovasc. Imaging *40*, 1221–1233.
18. Edwards, N.C., Moody, W.E., Yuan, M., Hayer, M.K., Ferro, C.J., Townend, J.N., and Steeds, R.P. (2015). Diffuse interstitial fibrosis and myocardial dysfunction in early chronic kidney disease. Am. J. Cardiol. *115*, 1311–1317.
19. Schuett, K., Marx, N., and Lehrke, M. (2023). The cardio-kidney patient: Epidemiology, clinical characteristics and therapy. Circ. Res. *132*, 902–914.
20. Vrudhula, A., Vukadinovic, M., Haeffele, C., Kwan, A.C., Berman, D., Liang, D., Siegel, R., Cheng, S., and Ouyang, D. (2025). Automated deep learning phenotyping of tricuspid regurgitation in echocardiography. JAMA Cardiol. https://doi.org/10.1001/jamacardio.2025.0498.
21. Vrudhula, A., Duffy, G., Vukadinovic, M., Liang, D., Cheng, S., and Ouyang, D. (2024). High-throughput deep learning detection of mitral regurgitation. Circulation *150*, 923–933.
22. Long, A., Finer, J., Hartman, H., Hartzel, D., Jing, L., Kelsey, C., Rocha, D., Ruhl, J., vanMaanen, D., Elnabawi, Y., et al. (2025). Deep learning for echocardiographic assessment and risk stratification of aortic, mitral, and tricuspid regurgitation: the DELINEATE-Regurgitation study. Eur. Heart J. https://doi.org/10.1093/eurheartj/ehaf248.
23. Holste, G., Oikonomou, E.K., Mortazavi, B.J., Coppi, A., Faridi, K.F., Miller, E.J., Forrest, J.K., McNamara, R.L., Ohno-Machado, L., Yuan, N., et al. (2023). Severe aortic stenosis detection by deep learning applied to echocardiography. Eur. Heart J. *44*, 4592–4604.
24. Duffy, G., Cheng, P.P., Yuan, N., He, B., Kwan, A.C., Shun-Shin, M.J., Alexander, K.M., Ebinger, J., Lungren, M.P., Rader, F., et al. (2022). High-Throughput Precision Phenotyping of Left Ventricular Hypertrophy With Cardiovascular Deep Learning. JAMA Cardiol *7*, 386–395.
25. Oikonomou, E.K., Sangha, V., Vasisht Shankar, S., Coppi, A., Krumholz, H.M., Nasir, K., Miller, E.J., Gallegos Kattan, C., Al-Mallah, M.H., Al-Kindi, S., et al. (2025). Artificial intelligence-enabled electrocardiography and echocardiography to track preclinical progression of transthyretin amyloid cardiomyopathy. Eur. Heart J. https://doi.org/10.1093/eurheartj/ehaf450.
26. Holste, G., Oikonomou, E.K., Tokodi, M., Kovács, A., Wang, Z., and Khera, R. (2025). Complete AI-enabled echocardiography interpretation with multitask deep learning. JAMA *334*, 306–318.
27. Goto, S., Solanki, D., John, J.E., Yagi, R., Homilius, M., Ichihara, G., Katsumata, Y., Gaggin, H.K., Itabashi, Y., MacRae, C.A., et al. (2022). Multinational federated learning approach to train ECG and echocardiogram models for hypertrophic cardiomyopathy detection. Circulation *146*, 755–769.

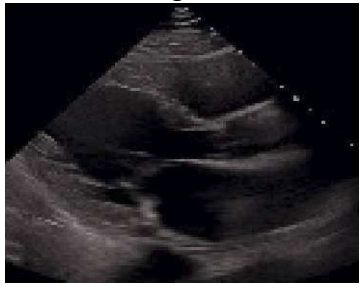
28. Vemulapalli, S., Alenezi, F., Jeong, H., Giczewska, A., Chiswell, K., Douglas, P., Zhang, A.R., Shah, S., Henao, R., and Wang, A. (2025). Machine learning computer vision point of care decision support of echocardiographic identification of hypertrophic cardiomyopathy. JACC Adv. *4*, 101746.
29. Sahashi, Y., Vukadinovic, M., Amrollahi, F., Trivedi, H., Rhee, J., Chen, J., Cheng, S., Ouyang, D., and Kwan, A.C. (2025). Opportunistic screening of chronic liver disease with deep-learning–enhanced echocardiography. NEJM AI *2*. https://doi.org/10.1056/aioa2400948.
30. Hughes, J.W., Yuan, N., He, B., Ouyang, J., Ebinger, J., Botting, P., Lee, J., Theurer, J., Tooley, J.E., Nieman, K., et al. (2021). Deep learning evaluation of biomarkers from echocardiogram videos. EBioMedicine *73*, 103613.
31. Kidney disease surveillance system https://wwwn.cdc.gov/kdss/Calculators.aspx.
32. Samad, Z., Sivak, J.A., Phelan, M., Schulte, P.J., Patel, U., and Velazquez, E.J. (2017). Prevalence and outcomes of left-sided valvular heart disease associated with chronic kidney disease. J. Am. Heart Assoc. *6*. https://doi.org/10.1161/JAHA.117.006044.
33. Ureña-Torres, P., D'Marco, L., Raggi, P., García-Moll, X., Brandenburg, V., Mazzaferro, S., Lieber, A., Guirado, L., and Bover, J. (2020). Valvular heart disease and calcification in CKD: more common than appreciated. Nephrol. Dial. Transplant *35*, 2046–2053.
34. Kipourou, K., O'Driscoll, J.M., and Sharma, R. (2022). Valvular heart disease in patients with chronic kidney disease. Eur. Cardiol. *17*, e02.
35. Halder, R.K., Uddin, M.N., Uddin, M.A., Aryal, S., Saha, S., Hossen, R., Ahmed, S., Rony, M.A.T., and Akter, M.F. (2024). ML-CKDP: Machine learning-based chronic kidney disease prediction with smart web application. J. Pathol. Inform. *15*, 100371.
36. Islam, M.A., Majumder, M.Z.H., and Hussein, M.A. (2023). Chronic kidney disease prediction based on machine learning algorithms. J. Pathol. Inform. *14*, 100189.
37. Debal, D.A., and Sitote, T.M. (2022). Chronic kidney disease prediction using machine learning techniques. J. Big Data *9*. https://doi.org/10.1186/s40537-022-00657-5.
38. Isaza-Ruget, M.A., Yomayusa, N., González, C.A., H, C.A., de Oro, F.A., V., Cely, A., Murcia, J., Gonzalez-Velez, A., Robayo, A., Colmenares-Mejía, C.C., et al. (2024). Predicting chronic kidney disease progression with artificial intelligence. BMC Nephrol. *25*, 148.
39. Bai, Q., Su, C., Tang, W., and Li, Y. (2022). Machine learning to predict end stage kidney disease in chronic kidney disease. Sci. Rep. *12*, 8377.
40. Holmstrom, L., Christensen, M., Yuan, N., Weston Hughes, J., Theurer, J., Jujjavarapu, M., Fatehi, P., Kwan, A., Sandhu, R.K., Ebinger, J., et al. (2023). Deep learning-based electrocardiographic screening for chronic kidney disease. Commun. Med. (Lond.) *3*, 73.
41. Kula, A.J., Prince, D.K., Katz, R., and Bansal, N. (2023). Mortality burden and life-years lost across the age spectrum for adults living with CKD. Kidney360 *4*, 615–621.
42. Segal, Z., Kalifa, D., Radinsky, K., Ehrenberg, B., Elad, G., Maor, G., Lewis, M., Tibi, M., Korn, L., and Koren, G. (2020). Machine learning algorithm for early detection of end-stage renal disease. BMC Nephrol. *21*, 518.
43. Kadappu, K.K., Abhayaratna, K., Boyd, A., French, J.K., Xuan, W., Abhayaratna, W., and Thomas, L. (2016). Independent echocardiographic markers of cardiovascular involvement in chronic kidney disease: The value of left atrial function and volume. J. Am. Soc. Echocardiogr. *29*, 359–367.
44. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2016). Grad-CAM: Visual explanations from deep networks via Gradient-based localization. arXiv [cs.CV].

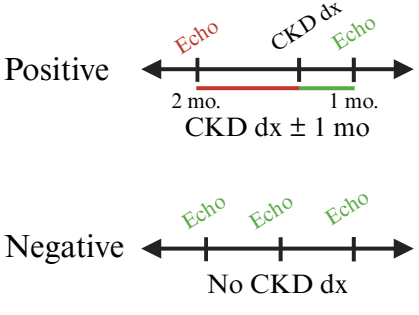
45. Hsu, C.-Y., Yang, W., Parikh, R.V., Anderson, A.H., Chen, T.K., Cohen, D.L., He, J., Mohanty, M.J., Lash, J.P., Mills, K.T., et al. (2021). Race, genetic ancestry, and estimating kidney function in CKD. N. Engl. J. Med. *385*, 1750–1760.
46. Tran, D., Wang, H., Torresani, L., Ray, J., LeCun, Y., and Paluri, M. (2017). A Closer Look at Spatiotemporal Convolutions for Action Recognition. arXiv [cs.CV].

Deep Learning for CKD Detection

Model Input

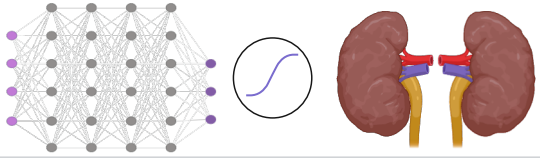
325,377 PLAX videos from 62,818 patients



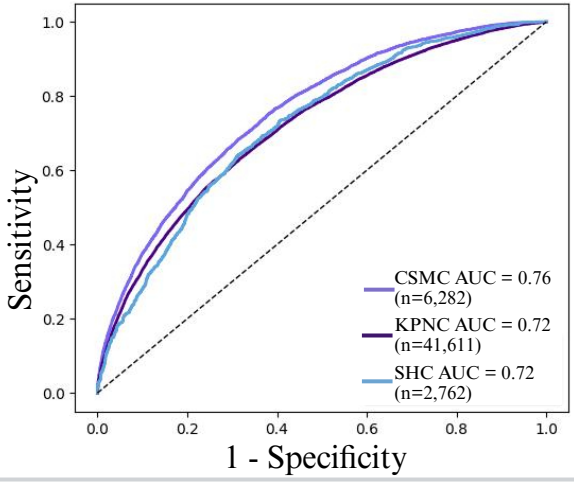


Model Output

Probability that CKD is present



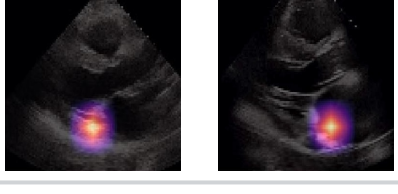
Model Performance



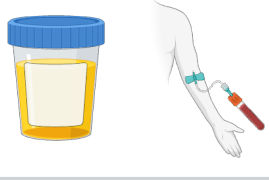
Clinical Utility of AI for CKD Screening

PLAX

Explainable AI Model



Clinical Management



Model Explainability

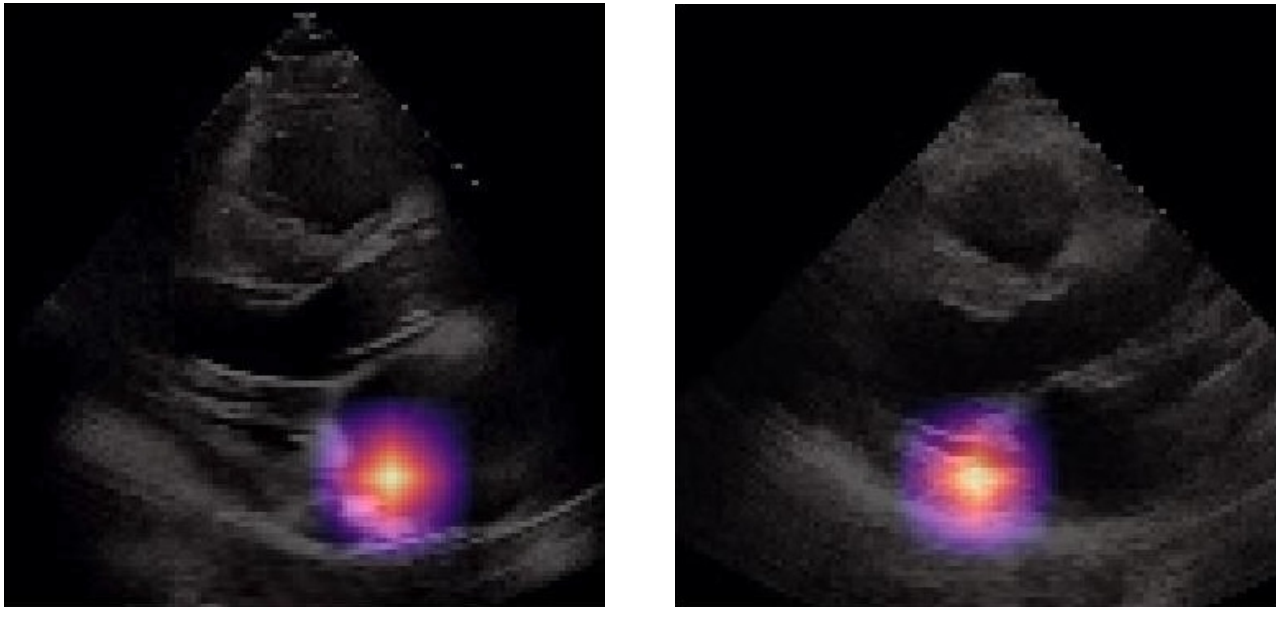


Table 1. Demographics and co-morbidities of the internal development cohort at Cedars-Sinai Medical Center (CSMC), the external validation cohort at Kaiser-Permanente Northern California (KPNC), and the external validation cohort at Stanford Healthcare (SHC)

	CSMC Controls (n=52,346)	**CSMC Cases (n=10,472)**	**KPNC Controls (n=31,357)**	**KPNC Cases (n=10,254)**	**SHC Controls (n=1,964)**	**SHC Cases (n=798)**
Age	67.7 ± 15.6	66.5 ± 15.2	69.2 ± 12.9	73.1 ± 13.1	61.0 ± 17.4	66.3 ± 14.9
Sex						
Male	33,645 (64.3%)	6,731 (64.3%)	15,763 (50.3%)	5,621 (54.8%)	876 (44.6%)	497 (62.3%)
Female	18,701 (35.7%)	3,741 (35.7%)	15,572 (49.7%)	4,631 (45.2%)	1088 (55.4%)	301(37.7%)
Race and Ethnicity						
American Indian or Alaska Native	106 (0.2%)	28 (0.3%)	154 (0.5%)	62 (0.6%)	7 (0.4%)	2 (0.3%)
Asian	3,594 (7.0%)	895 (8.6%)	4,695 (15.0%)	1,805 (17.6%)	384 (19.6%)	119 (14.9%)
Black or African American	6,072 (11.8%)	2,238 (21.4%)	2,191 (7.0%)	1,504 (14.7%)	103 (5.2%)	54 (6.8%)
Native Hawaiian/Pacific Islander	142 (0.3%)	46 (0.4%)	219 (0.7%)	167 (1.6%)	32 (1.6%)	23 (2.9%)
Hispanic	5,738 (11.2%)	1,715 (16.4%)	2,923 (9.3%)	1,154 (11.3%)	198 (10.1%)	111 (13.9%)
Non-Hispanic	44,390 (86.3%)	8,652 (82.6%)	28,434 (90.7%)	9,100 (88.7%)	1,635 (83.2%)	672 (84.2%)
White	36,612 (71.1%)	6,484 (61.9%)	17,152 (54.7%)	5,454 (53.2%)	1,021 (52.0%)	437 (54.8%)
Other	3,702 (7.2%)	664 (6.3%)	185 (0.5%)	37 (0.3%)	309 (15.7%)	143 (17.9%)
Unknown	2,118 (4.0%)	98 (0.9%)	191 (0.6%)	14 (0.1%)	108 (5.5%)	20 (2.5%)

Left ventricular ejection fraction %	52.8% ± 17.4%	56.9% ± 14.9%	57.9% ± 10.2%	53.5% ± 13.4%	57.9% ± 11.0%	51.0% ± 15.4%
Left ventricular end-diastolic volume index (mL/m^2)	46.6 ± 25.8	51.9 ± 27.8	49.7 ± 17.1	55.7 ± 23.8	45.8 ± 18.1	55.7 ± 30.4
Left ventricular end-systolic volume index (mL/m^2)	21.5 ± 17.4	25.9 ± 19.7	21.7 ± 12.7	27.7 ± 20.2	20.3 ± 14.1	30.3 ± 27.5
Estimated glomerular filtration rate (eGFR, $mL/min/1.73m^2$)	74.02 ± 26.6	33.3 ± 22.5	58.5 ± 5.7	38.1 ± 18.2	89.4 ± 21.0	57.1 ± 28.5
Urine albumin-to-creatinine ratio (uACR)	154.3 ± 492.2	379.3 ± 737.1	46.0 ± 86.4	129.2 ± 131.7	N/A	N/A
Coronary artery disease	12,317 (23.5%)	4,013 (38.3%)	13,497 (43.0%)	6,902 (67.3%)	407 (20.7%)	403 (50.5%)
Hypertension	22,572 (43.1%)	8,334 (79.6%)	23,436 (74.7%)	10,009 (97.6%)	858 (43.7%)	534 (66.9%)
Diabetes mellitus	9,058 (17.3%)	5,089 (48.6%)	9,579 (30.5%)	7,092 (69.1%)	477 (24.3%)	424 (53.1%)
Dyslipidemia	16,739 (32.0%)	5,171 (49.4%)	20,857 (66.5%)	9,021 (88.0%)	891 (45.4%)	593 (74.3%)

Table 2. Model performance on patients with laboratory evidence of CKD

eGFR threshold (mL/min/1.73 m^2)	**Cases**	**Controls**	**AUC**	**Sensitivity**	**Specificity**
30	966	4,761	0.741 (0.740, 0.743)	82.5% (82.3%, 82.7%)	48.8% (48.8%, 48.8%)
45	1,649	4,078	0.726 (0.725, 0.727)	77.1% (76.9%, 77.2%)	53.8% (53.7%, 53.9%)
60	2,503	3,224	0.733 (0.732, 0.734)	74.6% (74.5%, 74.7%)	58.4% (58.3%, 58.5%)

Table 3. Performance of model in discriminating any stage of CKD by patient subgroup in the CSMC internal test cohort with 95% confidence intervals in parentheses.

Subgroup	Number of patients	CKD Prevalence	AUC	Sensitivity	Specificity
Diabetes mellitus	1,438	508 (35.3%)	0.728 (0.716, 0.739)	77.4% (76.2%, 79.3%)	52.8% (50.8%, 53.4%)
Hypertension	3,086	835 (27.1%)	0.752 (0.743, 0.759)	75.0% (74.2%, 76.8%)	60.9% (59.4%, 61.1%)
Coronary artery disease	1,639	395 (24.1%)	0.756 (0.745, 0.767)	76.2% (74.9%, 78.4%)	59.8% (58.1%, 60.2%)
Dyslipidemia	2,195	500 (22.8%)	0.752 (0.742, 0.761)	72.8% (71.7%, 75.0%)	63.1% (61.7%, 63.5%)
EF < 50%	1,278	444 (34.7%)	0.700 (0.685, 0.715)	79.9% (77.9%, 81.9%)	45.6% (44.2%, 47.0%)
EF ≥ 50%	4,976	717 (14.4%)	0.770 (0.762, 0.779)	72.9% (71.3%, 74.5%)	66.9% (66.3%, 67.6%)
Male	4,043	652 (16.1%)	0.752 (0.743, 0.760)	75.4% (73.9%, 76.8%)	60.5% (59.7%, 61.2%)
Female	2,239	361 (16.1%)	0.762 (0.750, 0.775)	74.8% (72.7%, 76.9%)	64.0% (63.1%, 64.9%)
Age < 65 years	2,450	361 (14.7%)	0.820 (0.810, 0.830)	79.9% (78.2%, 81.6%)	68.0% (67.1%, 68.9%)
Age ≥ 65 years	3,948	652 (16.5%)	0.708 (0.700, 0.717)	72.1% (70.4%, 73.7%)	57.7% (56.9%, 58.4%)

Table 4. Performance of model in discriminating any stage of CKD by patient subgroup in the KPNC external validation cohort with 95% confidence intervals in parentheses.

Subgroup	Number of patients	CKD Prevalence	AUC	Sensitivity	Specificity
Diabetes mellitus	16,671	7,092 (42.5%)	0.684 (0.678, 0.691)	81.1% (80.4%, 81.8%)	40.8% (40.0%, 41.5%)
Hypertension	33,445	10,009 (29.9%)	0.691 (0.687, 0.696)	78.1% (77.4%, 78.7%)	46.4% (45.9%, 46.9%)
Coronary artery disease	20,399	6,902 (33.8%)	0.698 (0.693, 0.704)	80.9% (80.2%, 81.6%)	43.6% (42.9%, 44.2%)
Dyslipidemia	29,878	9,021 (30.2%)	0.706 (0.701, 0.711)	78.6% (77.9%, 79.2%)	48.1% (47.6%, 48.7%)
EF < 50%	1,915	711 (37.1%)	0.659 (0.635, 0.675)	86.4% (83.6%, 87.7%)	30.5% (29.5%, 33.8%)
EF ≥ 50%	8,431	1,682 (20.0%)	0.716 (0.706, 0.729)	74.4% (72.5%, 76.0%)	54.8% (54.1%, 56.1%)
Male	21,384	5,625 (26.3%)	0.713 (0.707, 0.719)	78.6% (77.8%, 79.4%)	50.0% (49.3%, 50.6%)
Female	20,203	4,634 (23.0%)	0.725 (0.719, 0.732)	75.8% (74.8%, 76.8%)	55.1% (55.5%, 55.7%)
Age < 65 years	11,947	2,207 (18.5%)	0.786 (0.777, 0.794)	77.8% (76.5%, 79.2%)	62.6% (61.8%, 63.4%)
Age ≥ 65 years	29,742	8,078 (27.2%)	0.690 (0.684, 0.695)	77.2% (76.5%, 77.9%)	47.9% (47.4%, 48.5%)

Table 5. Performance of model in discriminating any stage of CKD by patient subgroup in the SHC external validation cohort with 95% confidence intervals in parentheses.

Subgroup	Number of patients	CKD Prevalence	AUC	Sensitivity	Specificity
Diabetes mellitus	901	424 (47.1%)	0.704 (0.669, 0.741)	68.9% (64.1%, 73.8%)	63.6% (59.2%, 67.9%)
Hypertension	1,392	534 (38.4%)	0.683 (0.654, 0.712)	64.8% (60.4%, 69.2%)	61.1% (57.9%, 64.4%)
Coronary artery disease	810	403 (49.8%)	0.638 (0.595, 0.680)	63.1% (57.5%, 68.8%)	54.8% (49.6%, 69.8%)
Dyslipidemia	1,484	593 (40.0%)	0.695 (0.667, 0.723)	63.3% (59.0%, 67.6%)	63.8% (60.7%, 66.9%)
EF < 50%	1,047	419 (40.0%)	0.683 (0.646, 0.720)	79.2% (75.0%, 83.3%)	45.9% (41.0%, 50.8%)
EF ≥ 50%	1,756	404 (23.0%)	0.695 (0.672, 0.718)	60.8% (57.1%, 64.6%)	67.6% (65.5%, 69.6%)
Male	1,373	497 (36.2%)	0.687 (0.656, 0.717)	60.1% (55.2%, 64.9%)	66.2% (62.9%, 69.3%)
Female	1,389	301 (21.7%)	0.695 (0.660, 0.730)	61.9% (55.8, 67.8%)	68.6% (66.0%, 71.2%)
Age < 65 years	1,352	317 (23.4%)	0.745 (0.715, 0.776)	62.4% (56.6%, 68.1%)	74.3% (71.6%, 76.9%)
Age ≥ 65 years	1,411	481 (34.1%)	0.641 (0.608, 0.673)	59.6% (54.6%, 64.7%)	60.6% (57.5%, 63.6%)

49 Spadina Ave. Suite 200
Toronto ON M5V 2J1 Canada
www.biorender.com

Confirmation of Publication and Licensing Rights - Open Access

July 25th, 2026

Subscription Type:	*Institution - Academic*
Agreement number:	*EY2A02WM8D*
Publisher Name:	*Cell Med*
Figure Title:	*Figure 1*
Citation to Use:	*Created in BioRender. Yuan, V. (2026) https://BioRender.com/m6w0sz2*

To whom this may concern,

This document ("Confirmation") hereby confirms that Science Suite Inc. dba BioRender ("BioRender") has granted the following BioRender user: Victoria Yuan ("User") a BioRender Academic Publication License in accordance with BioRender's Terms of Service and Academic License Terms ("License Terms") to permit such User to do the following on the condition that all requirements in this Confirmation are met:

1) publish their Completed Graphics created in the BioRender Services containing both User Content and BioRender Content (as both are defined in the License Terms) in publications (journals, textbooks, websites, etc.); and
2) sublicense such Completed Graphics under "open access" publication sublicensing models such as CC-BY 4.0 and more restrictive models, so long as the conditions set forth herein are fully met.

Requirements of User:

1) All Completed Graphics to be published in any publication (journals, textbooks, websites, etc.) must be accompanied by the following citation either as a caption, footnote or reference for each figure that includes a Completed Graphic:
"Created in BioRender. Yuan, V. (2026) https://BioRender.com/m6w0sz2 ".
2) All terms of the License Terms including all Prohibited Uses are fully complied with. E.g. For Academic License Users, no commercial uses (beyond publication in journals, textbooks or websites) are permitted without obtaining or switching to a BioRender Industry Plan.
3) A Reader (defined below) may request that the User allow their figure to be a public template for Readers to view, copy, and modify the figure. It is up to the User to determine what level of access to grant.

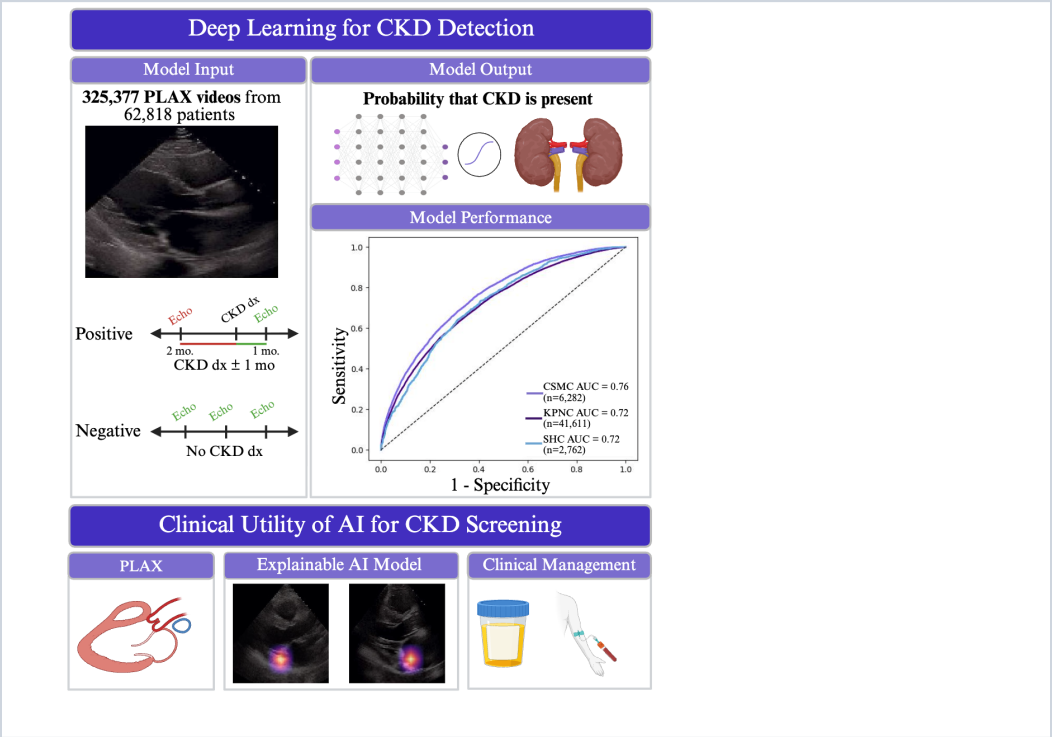
Open-Access Journal Readers:

Open-Access journal readers ("Reader") who wish to view and/or re-use a particular Completed Graphic in an Open-Access journal subject to CC-BY sublicensing may do so by clicking on the URL link in the applicable citation for the subject Completed Graphic.

The re-use/modification options below are available after the Reader requests the User to adapt their

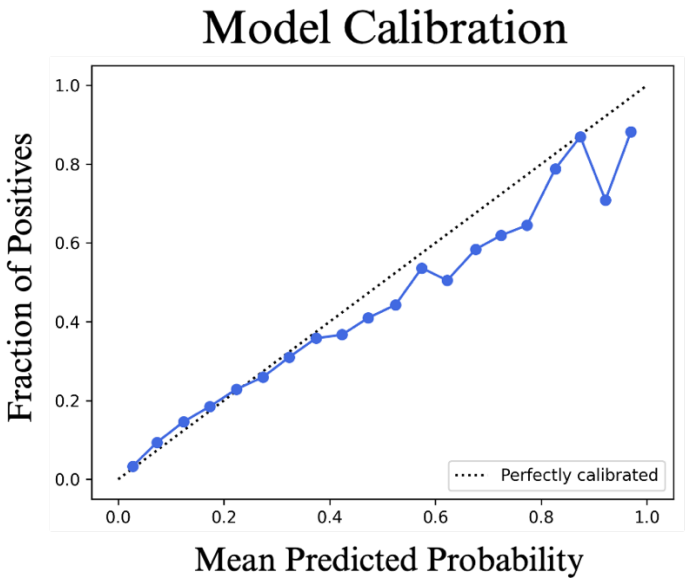
figure as a BioRender template and the User has granted such access.

1) View-Only/Free Plan Use: A Reader who wishes to only view the Completed Graphic may do so in the BioRender Services as either a BioRender Free Plan user or simply as a viewer. By becoming a BioRender Free Plan user, the Reader may view, modify and re-use the Completed Graphic as permitted under BioRender's Basic License Terms (e.g. personal use only, no publishing or commercial use permitted).
2) Re-Use/Publish with No Modifications: For any re-use and re-publication of a Completed Graphic with no modification(s) to the Completed Graphic made by the Reader, a Reader may do so by citing the original author using the citation noted above with the Completed Graphic. The Reader must also comply with the underlying License Terms which apply to the Completed Graphic as noted above (e.g. no commercial use for Academic License).
3) Re-Use/Publish with Modifications: For any re-use and re-publication of a Completed Graphic with a modification(s) made by the Reader, the Reader may do so by becoming a BioRender user themselves under either an Academic or Industry Plan, citing the original author using the citation noted above with the Completed Graphic and complying with the applicable License Terms.

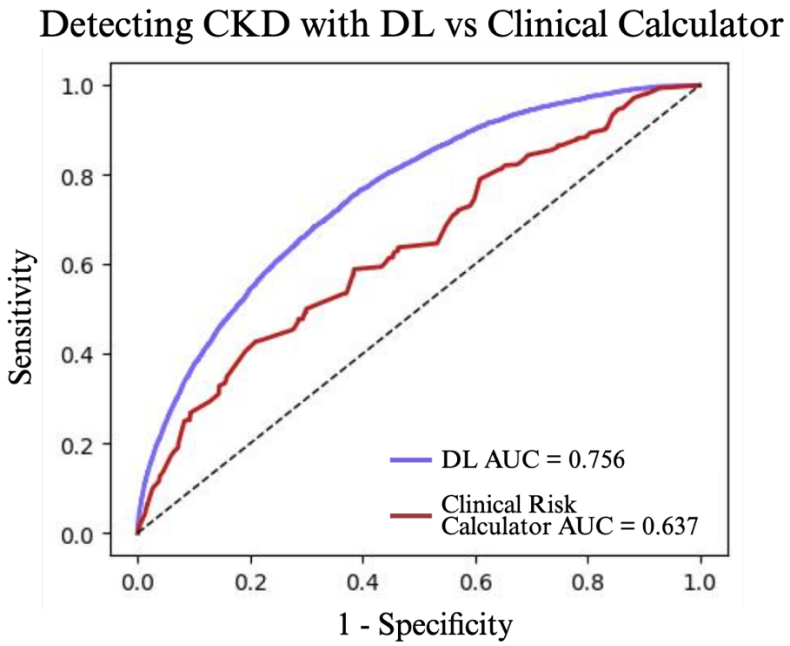


For any questions regarding this document, or other questions about publishing with BioRender, please refer to our BioRender Publication Guide, or contact BioRender Support at support@biorender.com.

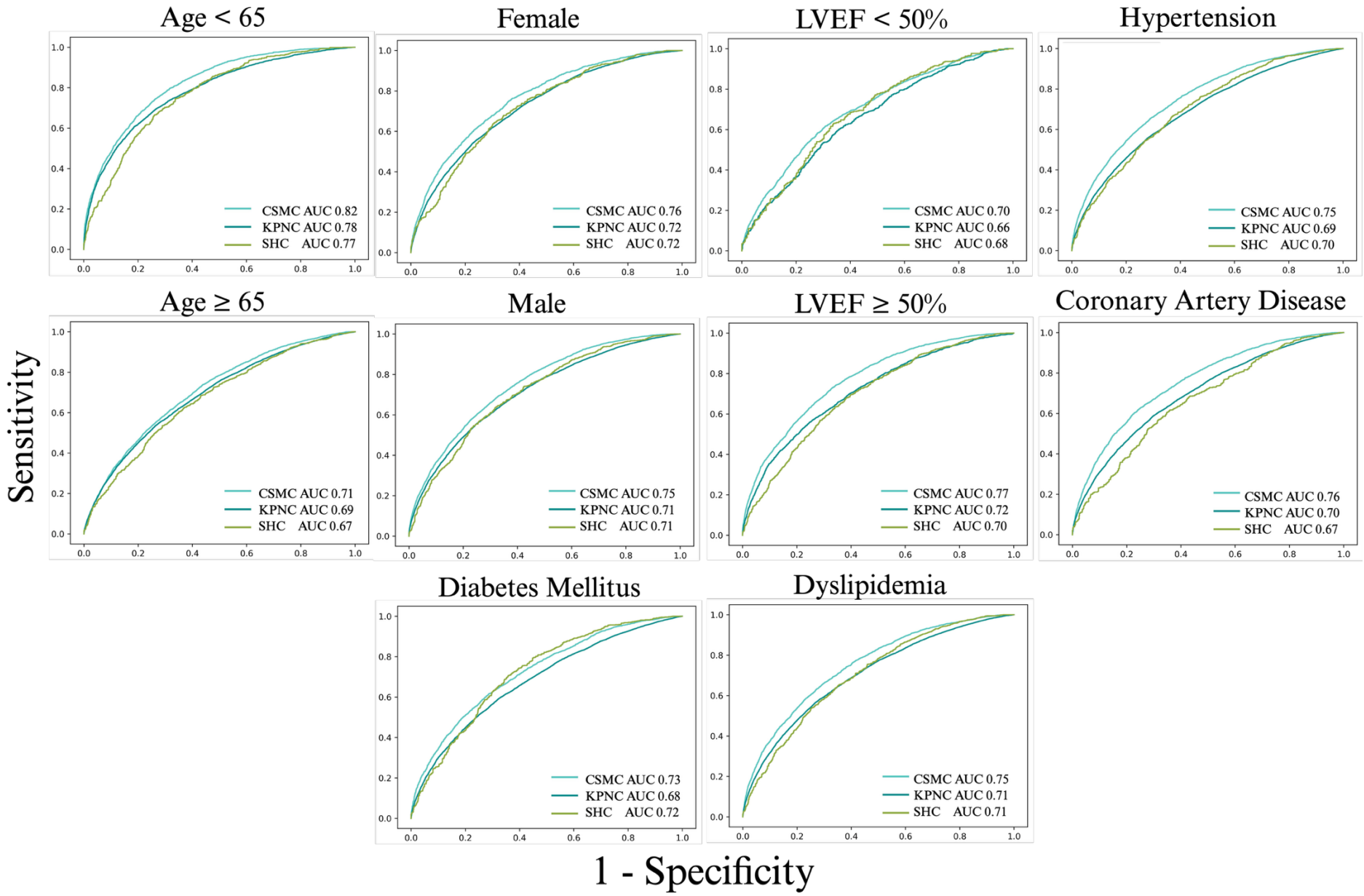
Supplementary Figure 1. Model calibration with calibration slope 1.15, calibration intercept -0.02, and Brier score loss of 0.11. Related to Figure 1.



Supplementary Figure 2. AUC to detect CKD using the clinical risk calculator recommended by the Center for Disease Control (red) and Prevention versus DL-based detection (purple). Related to Figure 1.



Supplementary Figure 3. Model performance in patient subgroups based on diabetes mellitus, hypertension, dyslipidemia, coronary artery disease, ejection fraction, age, and sex in the CSMC internal test and SHC and KPNC external validation cohorts. Related to Tables 3, 4, and 5.


Age < 65
CSMC AUC 0.82
KPNC AUC 0.78
SHC AUC 0.77
Female
CSMC AUC 0.76
KPNC AUC 0.72
SHC AUC 0.72
LVEF < 50%
CSMC AUC 0.70
KPNC AUC 0.66
SHC AUC 0.68
Hypertension
CSMC AUC 0.75
KPNC AUC 0.69
SHC AUC 0.70
Age ≥ 65
CSMC AUC 0.71
KPNC AUC 0.69
SHC AUC 0.67
Male
CSMC AUC 0.75
KPNC AUC 0.71
SHC AUC 0.71
LVEF ≥ 50%
CSMC AUC 0.77
KPNC AUC 0.72
SHC AUC 0.70
Coronary Artery Disease
CSMC AUC 0.76
KPNC AUC 0.70
SHC AUC 0.67
Diabetes Mellitus
CSMC AUC 0.73
KPNC AUC 0.68
SHC AUC 0.72
Dyslipidemia
CSMC AUC 0.75
KPNC AUC 0.71
SHC AUC 0.71
Sensitivity
1 - Specificity

Supplementary Table 1. ICD-10 codes used to identify patients diagnosed with CKD. Related to Figure 1 and Table 1.

ICD 10 Codes	N18.1 N18.2 N18.30, N18.31, N18.32 N18.4 N18.5 N18.6 N18.9

Supplementary Table 2. Demographics and co-morbidities of the patients in the train, validation, and test cohorts at CSMC for model development. Related to Table 1.

	Train (n=50,254)	**Validation (n=6,282)**	**Test (n=6,282)**
Age	67.3 ± 15.7	66.8 ± 16.0	67.0 ± 15.2
Sex			
Male	32,312 (64.3%)	4,021 (64.0%)	4,043 (64.4%)
Female	17,942 (35.7%)	2,261 (36.0%)	2,239 (35.6%)
Race and Ethnicity			
American Indian or Alaska Native	109 (0.2%)	10 (0.2%)	15 (0.2%)
Asian	3,562 (7.1%)	448 (7.1%)	479 (7.6%)
Black or African American	6,663 (13.3%)	820 (13.1%)	827 (13.2%)
Native Hawaiian/Pacific Islander	144 (0.3%)	17 (0.3%)	27 (0.4%)
Hispanic	5,948 (11.8%)	721 (1.5%)	784 (12.5%)
Non-Hispanic	42,430 (84.4%)	5,348 (85.1%)	5,264 (83.8%)
White	34,518 (68.7%)	4,358 (69.4%)	4,220 (67.2%)
Other	3,464 (6.9%)	418 (6.7%)	484 (7.7%)
Unknown	1,258 (2.5%)	123 (2.0%)	131 (2.1%)
Left ventricular ejection fraction %	56.3% ± 15.4%	56.7% ± 15.2%	56.2% ± 15.5%
Left ventricular end-diastolic volume index (mL/m^2)	47.5 ± 26.7	46.8 ± 23.1	47.6 ± 25.1
Left ventricular end-systolic volume index (mL/m^2)	22.1 ± 17.8	22.0 ± 18.6	22.3 ± 17.3
Estimated glomerular filtration rate (eGFR)	66.3 ± 30.4	66.1 ± 31.2	65.6 ± 30.7
Urine albumin-to-creatinine ratio (UACR)	137.0 ± 309.5	159.6 ± 337.3	165.4 ± 354.2
Coronary artery disease	13,015 (25.9%)	1,676 (26.7%)	1,639 (26.1%)
Hypertension	24,736 (49.2%)	3,084 (49.1%)	3,086 (49.1%)
Diabetes mellitus	11,290 (22.5%)	1,419 (22.6%)	1,438 (22.9%)

Dyslipidemia	17.503 (34.8%)	2,212 (35.2%)	2,195 (34.9%)

Supplementary Table 3. Model performance in discriminating different degrees of CKD in the CSMC cohort (top), KPNC cohort (middle), and SHC cohort (bottom). Negative examples were defined as PLAX videos without CKD. 95% confidence intervals are presented in parentheses. Related to Figure 1 and Table 2.

Cedars-Sinai Internal Test Cohort					
Degree of CKD	**Number of patients**	**ICD-10 Codes**	**AUC**	**Sensitivity**	**Specificity**
Mild	93	N18.1, N18.2	0.694 (0.670, 0.718)	68.6% (63.4%, 73.2%)	61.3% (61.3%, 62.4%)
Moderate	362	N18.3, N18.30, N18.31	0.701 (0.689, 0.712)	68.4% (65.5%, 70.1%)	61.3% (61.3%, 62.4%)
Severe	149	N18.4, N18.5	0.773 (0.755, 0.790)	76.6% (73.1%, 79.7%)	61.3% (61.3%, 62.4%)
ESRD	307	N18.6	0.840 (0.829, 0.850)	86.9% (84.6%, 88.3%)	61.3% (61.3%, 62.4%)
Kaiser-Permanente External Validation Cohort					
Degree of CKD	**Number of patients**	**ICD-10 Codes**	**AUC**	**Sensitivity**	**Specificity**
Mild	682	N18.1, N18.2	0.678 (0.661, 0.693)	72.2% (69.5%, 74.8%)	52.5% (52.0%, 52.9%)
Moderate	5,832	N18.3, N18.30, N18.31	0.679 (0.673, 0.684)	72.8% (71.9%, 73.7%)	52.5% (52.0%, 52.9%)
Severe	1,088	N18.4, N18.5	0.785 (0.773, 0.796)	84.8% (83.0%, 86.4%)	52.5% (52.0%, 52.9%)
ESRD	1,360	N18.6	0.830 (0.821, 0.839)	88.1% (86.8%, 89.4%)	52.5% (52.0%, 52.9%)
Stanford Healthcare External Validation Cohort					

Degree of CKD	Number of patients	ICD-10 Codes	AUC	Sensitivity	Specificity
Mild	337	N/A	0.686 (0.665, 0.707)	62.4% (58.9%, 65.8%)	65.1% (63.4%, 66.8%)
Moderate	420	N/A	0.742 (0.727, 0.757)	70.3% (67.8%, 72.6%)	65.1% (63.4%, 66.7%)
Severe	125	N/A	0.838 (0.819, 0.857)	85.4% (81.7%, 89.0%)	65.1% (63.4%, 66.7%)
ESRD	129	N/A	0.837 (0.818, 0.855)	82.3% (78.8%, 85.7%)	65.1% (63.4%, 66.8%)

Supplementary Table 4. AUC, sensitivity, and specificity of model performance with 95% confidence intervals in subgroups with equal prevalence of samples with and without the comorbidity of interest across all 3 clinical sites. Related to Figure 1.

Site	Comorbidity	AUC	Sensitivity	Specificity
CSMC	Diabetes	0.748 (0.722, 0.773)	76.2% (72.0%, 80.5%)	58.2% (56.0%, 60.4%)
	Hypertension	0.765 (0.741, 0.790)	77.0% (72.8%, 81.0%)	61.6% (59.4%, 63.7%)
	Dyslipidemia	0.745 (0.719, 0.770)	73.0% (68.7%, 77.3%)	61.5% (59.3%, 63.6%)
	CAD	0.726 (0.698, 0.752)	72.5% (68.0%, 76.9%)	60.2% (58.0%, 62.3%)
	Age ≥ 65	0.775 (0.751, 0.799)	77.5% (73.4%, 81.6%)	63.0% (60.8%, 65.1%)
	LVEF < 50%	0.731 (0.705, 0.760)	76.0% (71.7%, 80.1%)	56.9% (54.7%, 59.0%)
	Diabetes, hypertension, dyslipidemia	0.774 (0.749, 0.797)	79.7% (75.7%, 83.5%)	59.3% (57.1%, 61.4%)
KPNC	Diabetes	0.693 (0.664, 0.720)	73.8% (69.5%, 77.9%)	50.2% (48.0%, 52.4%)
	Hypertension	0.651 (0.620, 0.681)	60.8% (56.0%, 65.5%)	59.2% (57.1%, 61.4%)
	Dyslipidemia	0.704 (0.676, 0.732)	73.2% (69.2%, 77.7%)	54.7% (52.6%, 57.0%)
	CAD	0.703 (0.674, 0.730)	74.8% (70.5%, 78.9%)	52.5% (50.3%, 54.7%)
	Age ≥ 65	0.742 (0.715, 0.768)	79.0% (74.9%, 82.9%)	54.9% (52.7%, 57.1%)
	LVEF < 50%	0.723 (0.693, 0.749)	80.7% (76.8%, 84.6%)	46.5% (44.3%, 48.7%)
	Diabetes, hypertension, dyslipidemia	0.700 (0.673, 0.727)	70.8% (66.3%, 75.2%)	56.5% (54.3%, 58.7%)
SHC	Diabetes	0.705 (0.666, 0.744)	65.5% (58.8%, 72.0%)	64.9% (61.9%, 67.9%)
	Hypertension	0.703 (0.664, 0.742)	65.5% (58.9%, 72.0%)	65.0% (62.0%, 67.9%)
	Dyslipidemia	0.692	66.0%	64.2%

		(0.653, 0.729)	(59.2%, 72.4%)	(61.2%, 67.2%)
	CAD	0.701 (0.661, 0.739)	69.5% (62.9%, 75.7%)	59.9% (56.9%, 62.9%)
	Age ≥ 65	0.723 (0.685, 0.760)	67.0% (60.3%, 73.5%)	64.6% (61.7%, 67.5%)
	LVEF < 50%	0.714 (0.675, 0.752)	67.5% (61.0%, 73.9%)	62.5% (59.4%, 65.5%)
	Diabetes, hypertension, dyslipidemia	0.737 (0.684, 0.785)	68.0% (58.7%, 77.1%)	65.0% (60.8%, 69.1%)

Supplementary Table 5. AUC, sensitivity, and specificity with 95% confidence intervals for detection of moderate to severe CKD, excluding ESRD. Related to Figure 1 and Table 2.

Site	AUC	Sensitivity	Specificity
CSMC	0.721 (0.711, 0.731)	70.3% (68.4%, 72.1%)	61.9% (61.3%, 62.4%)
KPNC	0.695 (0.690, 0.701)	74.7% (73.9%, 75.5%)	52.5% (52.0%, 52.9%)
SHC	0.763 (0.750, 0.777)	73.8% (71.6%, 75.9%)	65.1% (63.4%, 66.7%)

Context and Significance

Chronic kidney disease (CKD) is the gradual loss of kidney function, impacting over 850 million people globally. CKD and cardiovascular disease share common risk factors, often occur together, and can lead to higher morbidity. Diagnosing CKD requires blood and urine studies, and over 60% of people with the condition do not know they have it. Researchers at Cedars-Sinai Medical Center, Kaiser Permanente Northern California, and Stanford University show that deep learning can reliably detect CKD from echocardiography, an ultrasound of the heart. Their deep learning model offers a noninvasive avenue to diagnose CKD earlier, prevent disease progression, and connect people to appropriate treatments that can protect both their heart and their kidney.

Highlights

- Deep learning can detect any stage of CKD from echocardiography.
- Our model demonstrates robust performance across 3 distinct clinical centers.
- Our model learns echocardiographic features of CKD shared across high-risk populations.

eTOC

Chronic kidney disease (CKD) has an underdiagnosis rate of over 60%, and high-risk populations are not routinely screened. Here, Yuan et al. develop and externally validate a deep learning model to detect CKD from echocardiography, offering a promising avenue to enable earlier screening and close diagnostic gaps.

CELL PRESS DECLARATION OF INTERESTS POLICY

Transparency is essential for a reader's trust in the scientific process and for the credibility of published articles. At Cell Press, we feel that disclosure of competing interests is a critical aspect of transparency. Therefore, we require a "declaration of interests" section in which all authors disclose any financial or other interests related to the submitted work that (1) could affect or have the perception of affecting the author's objectivity or (2) could influence or have the perception of influencing the content of the article.

What types of articles does this apply to?

We require that you disclose competing interests for all submitted content by completing and submitting the form below. We also require that you include a "declaration of interests" section in the text of all articles even if there are no interests to declare.

What should I disclose?

We require that you and all authors disclose any personal financial interests (e.g., stocks or shares in companies with interests related to the submitted work or consulting fees from companies that could have interests related to the work), professional affiliations, advisory positions, board memberships (including membership on a journal's advisory board when publishing in that journal), or patent applications and/or registrations that are related to the subject matter of the contribution. As a guideline, you need to declare an interest for (1) any affiliation associated with a payment or financial benefit exceeding $10,000 p.a. or 5% ownership of a company or (2) research funding by a company with related interests. You do not need to disclose diversified mutual funds, 401ks, or investment trusts.

Authors should also disclose relevant financial interests of immediate family members. Cell Press uses the Public Health Service definition of "immediate family member," which includes spouse and dependent children.

Where do I declare competing interests?

Competing interests should be disclosed on this form as well as in a "declaration of interests" section in the manuscript. This section should include financial or other competing interests as well as affiliations that are not included in the author list. Examples of "declaration of interests" language include:

> "AUTHOR is an employee and shareholder of COMPANY."
> "AUTHOR is a founder of COMPANY and a member of its scientific advisory board."

NOTE: Primary affiliations should be included with the author list and do not need to be included in the "declaration of interests" section. Funding sources should be included in the "acknowledgments" section and also do not need to be included in the "declaration of interests" section. (A small number of front-matter article types do not include an "acknowledgments" section. For these articles, reporting of funding sources is not required.)

What if there are no competing interests to declare?

If you have no competing interests to declare, please note that in the "declaration of interests" section with the following wording:

> "The authors declare no competing interests."

CELL PRESS DECLARATION OF INTERESTS FORM

If submitting materials via Editorial Manager, please complete this form and upload with your initial submission. Otherwise, please email as an attachment to the editor handling your manuscript.

Please complete each section of the form and insert any necessary "declaration of interests" statement in the text box at the end of the form. A matching statement should be included in a "declaration of interests" section in the manuscript.

Institutional affiliations

We require that you list the current institutional affiliations of all authors, including academic, corporate, and industrial, on the title page of the manuscript. ***Please select one of the following:***

☐ All affiliations are listed on the title page of the manuscript.

☐ I or other authors have additional affiliations that we have noted in the "declaration of interests" section of the manuscript and on this form below.

Funding sources

We require that you disclose all funding sources for the research described in this work. ***Please confirm the following:***

☐ All funding sources for this study are listed in the "acknowledgments" section of the manuscript.*

*A small number of front-matter article types do not include an "acknowledgments" section. For these, reporting funding sources is not required.

Competing financial interests

We require that authors disclose any financial interests and any such interests of immediate family members, including financial holdings, professional affiliations, advisory positions, board memberships, receipt of consulting fees, etc., that:

(1) could affect or have the perception of affecting the author's objectivity, *or*
(2) could influence or have the perception of influencing the content of the article.

Please select one of the following:

☐ We, the authors and our immediate family members, have no financial interests to declare.

☐ We, the authors, have noted any financial interests in the "declaration of interests" section of the manuscript and on this form below, and we have noted interests of our immediate family members.

Advisory/management and consulting positions

We require that authors disclose any position, be it a member of a board or advisory committee or a paid consultant, that they have been involved with that is related to this study. We also require that members of our journal advisory boards disclose their position when publishing in that journal. ***Please select one of the following:***

☐ We, the authors and our immediate family members, have no positions to declare and are not members of the journal's advisory board.

☐ The authors and/or their immediate family members have management/advisory or consulting relationships noted in the "declaration of interests" section of the manuscript and on this form below.

Patents

We require that you disclose any patent applications and/or registrations related to this work by any of the authors or their institutions. ***Please select one of the following:***

☐ We, the authors and our immediate family members, have no related patent applications or registrations to declare.

☐ We, the authors, have a patent application and/or registration related to this work, which is noted in the "declaration of interests" section of the manuscript and on this form below, and we have noted the patents of immediate family members.

Please insert any "declaration of interests" statements in this space. This exact text should also be included in the "declaration of interests" section of the manuscript. If no authors have a competing interest, please insert the text, "The authors declare no competing interests."

☐ **On behalf of all authors, I declare that I have disclosed all competing interests related to this work. If any exist, they have been included in the "declaration of interests" section of the manuscript.**