Additive prognostic value of aortic arch calcification of chest X-ray and feasibility of machine learning algorithm on cardiovascular outcome; retrospective study

Seungkyo Park

Department of Medicine

The Graduate School, Yonsei University

Additive prognostic value of aortic arch calcification of chest X-ray and feasibility of machine learning algorithm on cardiovascular outcome; retrospective study

Directed by Professor Eui-Young Choi

The Master's Thesis
submitted to the Department of Medicine,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree of
Master of Medical Science

Seungkyo Park

June 2021

This certifies that the Master's Thesis of Seungkyo Park is approved.

Thesis Supervisor : Eui-Young Choi

Thesis Committee Member#1 : Sungjun Kim

Thesis Committee Member#2 : Chul Hwan Park

The Graduate School
Yonsei University

June 2021

ACKNOWLEDGEMENTS

I would first like to thank my thesis advisor Professor Eui-Young Choi of the Division of Cardiology, Department of Internal Medicine at Yonsei University College of Medicine. The door to Prof. Choi's office was always open whenever I ran into a trouble spot or had a question about my research or writing. He consistently allowed this paper to be my own work,but steered me in the right the direction whenever he thought I needed it.

I would also like to thank the experts who were involved in the validation survey for this research project: Professor Sungjun Kim, a chairman of the thesis committee member, for his valuable critique, and most especially, for promoting the researchers to pursue her Master degree. Professor Chul Hwan Park, Assistant Professor of Department of Radiology at Yonsei University College of Medicine for his inspiration and intellectual comments. Professor Sang Chul Hwang, for his dedication of technical advices in terms of developing deep learning algorithm. Without their passionate participation and input, the validation survey could not have been successfully conducted.

Finally, I must express my very profound gratitude to my parents, husband and lovely daughter for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.
Seungkyo Park

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Additive prognostic value of aortic arch calcification of chest x ray and feasibility of machine learning algorithm on cardiovascular outcome; retrospective study

Seungkyo Park

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Eui-Young Choi)

Cardiovascular death is one of most common cause of death in the world. Traditional cardiovascular prognosis is based on Framingham risk scoring system or ACC/AHA Pooled Cohort Equation. Though those are excellent and convenient scoring system, we cannot deny that traditional calculation is not intuitive. Many studies revealed the relationship between aortic calcification and major cardiovascular events. Aortic arch is one of the most vulnerable segment in aorta against arterial pressure and arteriosclerosis. Although most of us agree with that aortic arch calcification on simple chest x-ray would be valuable, owing to its low reproducibility and validity of semi-quantative grading system, research focusing on aortic arch calcification and simple chest x-ray has not rigorously performed yet. In this study, we examines clinical implication of aortic arch calcification on chest x-ray to cardio-cerebrovascular outcome. After revealing its clinical importance and usefulness, then we developed machine learning algorithm to grading of aortic arch calcification on simple chest x-ray. Study population were collected patients who underwent carotid Doppler ultrasound at Gangnam Severance Hospital from 2009

March to 2012 February. Till now, total 3,080 patients were reviewed. Among them, 2,273 patients were finally enrolled in the study. Aortic arch calcification grade on chest x-ray was significantly correlated with presence of carotid artery plaques and pulse wave velocity, which represents arterial stiffness. CVA and all-cause death were significantly associated with aortic arch calcification grade on chest x-ray. The rate of admission due to heart failure aggravation was also highly related in patients whose aortic arch calcification grade was 3. In contrast, treatment with PTCA or any composite CVE was neither associated with aortic arch calcification grade on chest x ray. Predictive value of aortic arch calcification grade was also notable in case of CVA, all-cause death and some cases of admission rate due to heart failure aggravation. Through this study, we recognized that arteriosclerosis contributes to aortic arch calcification and its mechanism of action is different from atherosclerosis of coronary artery. In addition, we can also assume that aortic arch calcification grade has additive predictive value in addition to FRS for cardiovascular outcome. Further study would be elaborated to deep learning algorithm we developed so that clinicians can be instantly warned for risk of cardio-cerebrovascular outcome when they checked chest x-ray of patients.

Key words:
aortic arch calcification, chest x ray, cardiovascular outcome, deep learning algorithm, arteriosclerosis, atherosclerosis

Additive prognostic value of aortic arch calcification of chest X-ray and feasibility of machine learning algorithm on cardiovascular outcome; retrospective study

Seungkyo Park

Department of Medicine
The Graduate School, Yonsei University
(Directed by Professor Eui-Young Choi)

I. INTRODUCTION

1. Research Background

A. Cardio-cerebrovascular event as a major cause of death

For the past ten years, domestic statistical indicators revealed how deaths from heart disease and cerebrovascular disease consistently rank as a top 10 cause of death [1]. According to a 2019 National Statistical Office publication, death rate per 100,000 population due to cardiovascular disease and cerebrovascular disease are 62.4 and 44.7, respectively. Furthermore, its proportion is gradually increasing. In 2018, cardio-cerebrovascular disease was the second leading cause of death in Korea and an increase in deaths due to cardio-cerebrovascular disease is an apparent trend. In

United States, cardio-cerebrovascular disease is the number one cause of death, followed by death by malignancy [2]. In the Korean context, cardio-cerebrovascular disease is predicted to be a major cause of death in the future due to tendencies to favor western lifestyles [3]. Research on the prognostic factors of cardio-cerebrovascular disease is meaningful as the diagnosis rate of cardiovascular disease is increasing not only in terms of mortality rates, but also due to an aging population and the prolongation of lifespan due to the advancements in medicine.

B. Methods of prediction of cardiovascular events: Summary of previous literatures

(1) Prediction of cardiovascular outcome

Framingham Risk Score (FRS)

Conventionally, the prognosis for cardiovascular disease is predicted using the scoring system and the Framingham scoring system (FRS). FRS was introduced in the third report (Adult Treatment Panel III) elucidated by the National Cholesterol Education Program Expert Work Group on Diagnosis, Evaluation and Treatment of High Blood Cholesterol in Adults [4]. FRS is a calculation method that predicts the prognosis by calculating the score by utilizing the underlying disease and individual demographic characteristics and dividing the probability of cardiovascular accidents occurrence into low/medium/high risk within a 10-year timeframe. According to the stratification of FRS risks, target values of low-density lipoprotein (LDL) and drug treatment policies can be determined [5].

Adult Treatment Panel (ATP III) hard CHD risk score

ATP III hard CHD risk score (2002) is also known as modified FRS. It contains treatment of hypertension, age-specific points for smoking and total cholesterol while withdrawal of diabetes because diabetes is supposed to be equivalent to CHD already. Modified FRS was developed due to possibility of overestimation of 10-year CHD risk; also, it was based on the database of 1960s and 1970s cohort Caucasian group.

ACC/AHA Pooled Cohort Equations; ACC/AHA Guideline on the Assessment of cardiovascular risk

Around 2008, the National Heart, Lung and Blood Institute (NHLBI) designed a study that occurs within 10 years on predictive prognostic factors for atherosclerotic cardiovascular disease (ASCVD) which are associated with atherosclerosis. The reasons for risk scoring with FRS alone was that the patient group was changed, and the representative was lowered, the difference in race was not reflected, and problems were raised that the specific description of the endpoint was insufficient.

The research team gathered several groups of other cohorts, including African Americans, not a Framingham cohort group that was studied in FRS, and followed up for several years, and then published a new method for calculating cardiovascular prognosis, which is different from FRS, under the name ACC/AHA Pooled Cohort Equation [6]. If the target LDL value is presented in accordance with the existing FRS score, the 2013 ACC/AHA guideline emphasized the treatment group rather than the target LDL value. Accordingly, the following population characteristics were considered as treatment groups.

- Individuals with clinical ASCVD
- Individuals with primary elevations of LDL≥190mg/dL
- Individuals aged 40 to 75, diabetes and LDL level with 70 to 189 mg/dL
- Individuals without clinical ASCVD, nor diabetes and who are aged of 40 to 75 with LDL level 70 to 189 mg/dL, and finally 10-year ASCVD risk using Pooled Cohort Equation ≥ 7.5%

  Here, ASCVD is defined as the following four, and the endpoint used in the ACC/AHA Pooled Cohort Equation is commonly referred to as “first hard ASCVD” in the literature [7].

- Non fatal myocardial infarction
- Coronary heart disease death

- Fatal stroke
- Non fatal stroke

The Pooled Cohort Equation is calculated based on the patient's gender, age, race, total cholesterol, HDL cholesterol, systolic blood pressure, diabetes and smoking history [8]. Similar calculation methods for the prevention of cardiovascular disease and similar parameters like that of the Pooled Cohort Equation are used by the Joint British Societies (JBS3) [9].

(2) Prediction of cerebrovascular outcome

Framingham Stroke Risk Score (FSRS) and Revised-Framingham Stroke Risk Score (R-FSRS)

Risk prediction in terms of cerebrovascular outcome also exists. Framingham Stroke Risk Stroke (FSRS) and its modified score system, Revised-Framingham Stroke Risk Score (R-FSRS) are representative prediction model [10]. Various studies are still performed to validate prediction ability of new or traditional prediction methods [10].

2. Aortic Arch Calcification as a tool of prediction for cardiovascular disease and prognosis

The aorta is a structure that not only acts as a simple conduit which supplies blood flow to the whole body, but also changes the stiffness of the artery due to proliferation of the intima, occurrence and progression of atherosclerosis, and proliferation of the medial membrane due to hypertension and old age, and fibrosis.

Calcification in the body can be thought of as divided into the calcification of bones and areas other than bones. The material that makes up the bone is hydroxyapatite, in addition to hydroxyapatite. Other than bones, the calcareous components include Magnesium whitlockite, calcium phosphate deposit, and dystrophic calcification. Calcification is also associated with cell necrosis. Among these, vascular calcification can be anatomically divided into arteriosclerosis and atherosclerosis. Atherosclerosis mainly forms a linear deposit in the medial layer among elderly, chronic renal failure, and diabetic patients. Atherosclerosis usually occurs in the intima layer and is a form of calcification which occurs in the atherosclerotic plaque, which has numerous implications. Atherosclerosis (arterial stiffness) can be regarded as a functional defect of arterial blood vessels due to atherosclerosis. Symptoms such as increased PWV, increased LV afterload, and decreased perfusion is eventually associated with the manifestation of hemodynamic compromise.

It is well known that arterial stiffness typically has a correlation with the degree of atherosclerosis in the body [11]. Arterial vascular calcification is generally known to represent arteriosclerosis. In particular, many studies conducted reveal how aortic calcification has a correlation with coronary vascular calcification. Various radiological equipment methods are used to measure calcification, among which CT is most commonly used. In addition, studies on aortic calcification using DEXA are actively being administered [12]. Research on aortic calcification using simple x-rays has also been published, but questions about the reproducibility and validity of the calcification degree grading has been raised [13]. This phenomenon is very unfortunate. By conducting a simple chest x-ray, the calcification of the aortic arch is confirmed instantaneously. The aortic arch is the most susceptible to atherosclerosis among the thoracic aorta, so calcification of the aortic arch is conducted via a simple x-ray. This is because the analysis of data on the cardiovascular system is predicted to be cost-effective for the study of prognostic factors of the cardiovascular system.

In regard to the simple chest x-ray of aortic arch calcification grading system, a semi-quantitative scoring system is typically used [13-16].

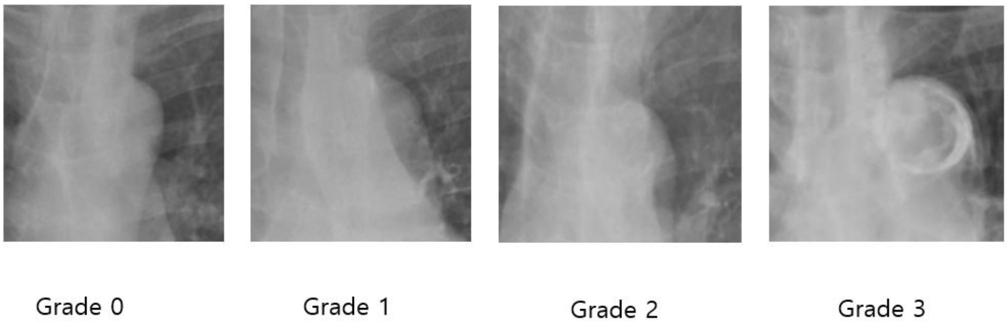


Figure 1. Grading of aortic arch calcification on simple chest x ray [14-16]

3. Machine Learning and its application to CXR

A CXR can be used to suspect lung field lesions, pleural hypertrophy, diaphragm position abnormalities, cardiac hypertrophy, mediastinal width, and structural abnormalities of the aorta and aortic arches. CXRs are largely taken in two ways. As posterior and anterior pictures are more accurate than before and after pictures, they are more useful for obtaining clinical clues and delineating interpretations.

As various medical imaging devices such as CT, MRI, etc. which confirm the cross section vertically and horizontally have developed and have increased in accessibility, clinicians' interest in basic imaging tests such as CXRs are comparatively decreasing. However, CXR is still one of the basic imaging tests performed not only by surgical patients, but also by those within the medical field.

The largest number of single imaging tests is CXR. Looking at the recent trend of CXR research through pubmed searches, nurses, radiologists, and various other professions attempt to read CXR images in addition to doctors [17, 18]. Studies have shown that said alternative professionals who read x-ray results are not inferior to that of doctors. In addition, studies related to the remote reading of CXRs using Google Glass technology are also being conducted [19]. Since 2012, scholars have begun to publish research results that combine artificial intelligence (AI) and machine learning into medical diagnostic tests and treatments. Chest x-ray and machine learning convergence research began in earnest from 2016, and the number has been increasing exponentially until 2020 (Table 1). As a result of searching pubmed for “Machine Learning, Medicine”, an effort to apply artificial intelligence systems to various medical fields can be seen in general medicine as published since 2010 by the Department of Radiology, Anesthesiology, Pathology, Oncology, Respiratory Medicine, Ophthalmology, Rehabilitation Medicine, and Renal Medicine.

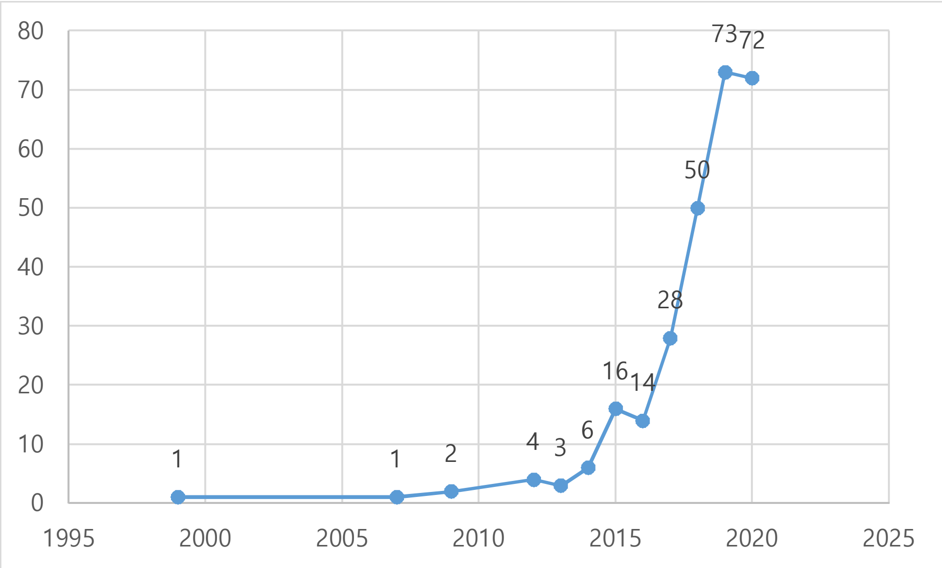


Figure 9 Number of articles searched with keywords "chest X-ray machine learning"

Research results that combine machine learning with simple chest x-ray images are being published within the field of imaging medicine. Results that can be obtained by analyzing simple chest X-rays through machine learning include pneumothorax [21], pneumonia, pulmonary pleural effusion, and atelectasis. However, there are no studies that study the significance of learning and the clinical application of aortic arch calcification through machine learning which can be applied as a prognostic factor for the cardiovascular system.

4. Research Objectives

In this study, through a retrospective approach, we investigate whether the degree of aortic arch calcification on a simple chest X-ray is related to cardio-cerebrovascular accidents through machine learning. In addition to the prognostic factors of vascular disease, we will investigate whether independent or additional prognostic information can be attained. The hypothesis that summarizes this research is as follows.

① The calcification of the aortic arch on a simple chest X-ray reflects the arterial stiffness proportionally.

② Aortic arch calcification on simple chest X-ray has an independent significance or additional prognostic factor for the cardiocerebrovascular outcome.

③ Grading of the aortic calcification on chest PA is possible with Deep Learning and this has the same effect as the prognostic factor of ②.

II. MATERIALS AND METHODS

This is a multidisciplinary research study conducted in the internal medicine department by the radiological science research institute and the radiology department. As the main pivot of the internal medicine, patients were recruited retrospectively, and after anonymizing simple chest X-rays of patients in collaboration with the Department of Radiology, a semi-quantitative grade is calculated with the naked eye. The anonymized image and the aortic arch calcification grade were gathered in order to complete the algorithm through machine learning at the Institute of Radiological Medicine. The algorithm was then applied to the patient group so that the aortic arch calcification on a simple chest X-ray is added to the prognostic factors of the cardiovascular system. By doing so, assessments are made to verify whether the aortic arch calcification plays an additional or exclusive role in terms of the outcome factor.

In order to complete the desired algorithm using Machine Learning, a training dataset for machine learning and a testing dataset to check the results of training after learning are required. Therefore, in this study, simple chest X-rays of patients, whom are the main subjects of the study, who underwent carotid doppler ultrasound in the echocardiography room from January 1, 2007 to March 31, 2012 were comprehensively different from the patient group before the application of machine learning. As such, it was determined that a patient group would be recruited and image training with a learning set would be conducted. This study was approved by institutional review board of Gangnam Severance Hospital (3-

2021-0051) and the informed consent was waived in the nature of retrospective study.

1. Target patient group (inclusion criteria)

① Training Dataset Collection

- Aortic arch calcification uses 500 images for each grade as a training set for machine learning.
- Grade 2 and 3 patient groups were difficult to observe in the general population, so they were extracted as follows using SCRAP 2.0, a big data software used at Gangnam Severance Hospital.
  - ■ When simple chest x-ray reading contains the content of “aortic arch calcification”:
    - ◆ From March 31, 2014 to September 31, 2020, we searched for cases in which aortic calcification words are included in the reading among patients who underwent simple chest X-rays at Gangnam Severance Hospital.
  - ■ Groups for Representative of aorta calcification
    - ◆ Hemodialysis unit
    - ◆ stroke unit
    - ◆ aorta surgery
- Grade 0, 1 patient groups are listed in the database of the echocardiography laboratory for patients who were tested in the echocardiography laboratory of Gangnam

Severance Hospital from July 1, 2020 to September 30, 2020. Out of patients who underwent carotid Doppler ultrasound in the echocardiography room of Gangnam Severance Hospital until March 31, 2012, patients who have been tested for simple chest posterior x-ray images about a year before or after performing carotid Doppler ultrasound.

② Testing Dataset Collection

Are the intended target.

2. Excluded patient group (exclusion criteria)

Patients who have not been tested for simple chest posterior x-ray images for one year before and after carotid Doppler ultrasound ①

① Patients unable to analyze simple chest posterior X-ray image due to poor quality.

② A patient who took an EKG monitor attached when taking a simple chest posterior anterior X-ray screening.

③ Patients who have an image that can cause confusion in machine learning due to graft or suture after undergoing surgery on the area, including the aortic arch when taking a simple chest posterior X-ray scan.

④ Unable to analyze the patient's result due to a poor-quality carotid doppler ultrasound image

⑤ The time difference between carotid Doppler ultrasound and chest x ray is greater than 3 years.

3. Screening technique: medical record check

4. Target number of subjects and basis of calculation: For patients who underwent carotid ultrasound and doppler in the echocardiography room of Gangnam Severance Hospital from January 1, 2007 to March 31, 2012. 5,034 people in the medical record.

3. chest x-ray acquisition

① The anonymized image is extracted by obtaining a list of patients in ①, anonymizing it, and assigning a new study number. The aortic arch calcification stage on simple chest radiography x-ray is graded 0, 1, 2, 3, and the validation of images make the training dataset data of machine learning. The training dataset goes through the validation phase again.

② Testing dataset: Acquire the list of patients in ② and go through the process of anonymizing the images of the training dataset. By making use of the developed machine learning algorithm, grading results outputs are obtained. These results are validated by radiologists and clinicians in order to check the outcomes and analyze whether they have statistical and clinical significance.

4. Development of deep learning algorithm; Automated aortic arch calcification grading for simple chest X-rays using deep learning

In order to automatically determine the grade of aortic arch calcification in simple X-ray images, we divided the study into two stages.

First, the region of interested calcification(ROIC) area was extracted from the chest radiographic image using the SegNet network. Next, the aortic arch calcification grade was automatically determined through the Alexnet network.

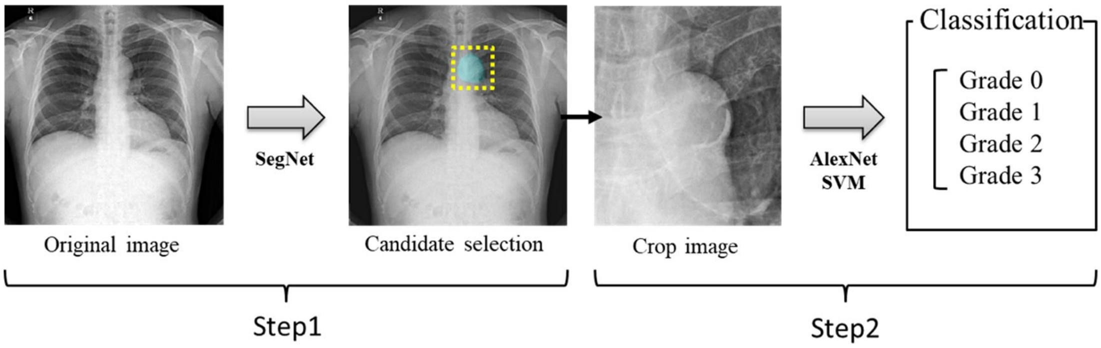


Figure 10. Workflow diagram for the classification

A. Study participants and radiograph data

A total of 2499 simple chest X-ray images were included in this study.

For methodological consistency, cases with poor image quality, which contain medical devices such as EKG leads, pacemaker and cases with history of surgery of aortic arch were excluded from the dataset used in the developed algorithm. Finally, 161 patients were used to extract the central position of the aortic arch, and 1087 patients were used to determine the calcification grade of the aortic arch (Fig 4).

Also, the data to extract the central position of the aortic arch were divided into training set (70%), validation set (15%) and test set. (15%). And the data for determining the aortic arch calcification grade were divided into training set (80%), validation set (10%), and test set (10%).

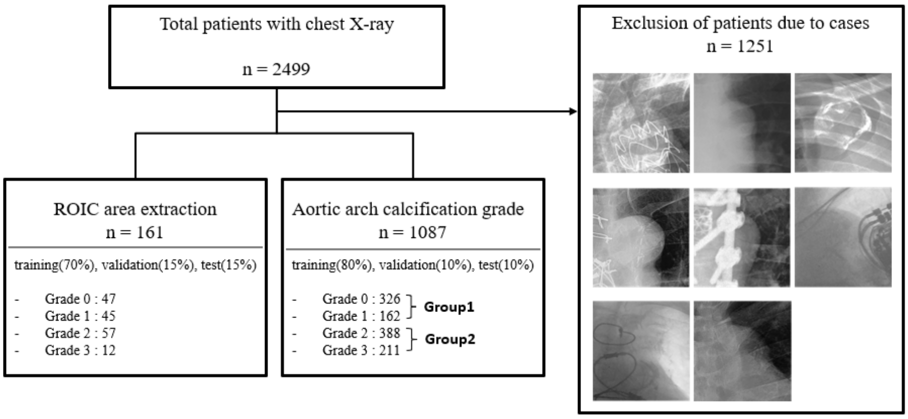


Figure 4. Indicate patient numbers after application of inclusion and exclusion criteria

B. ROIC extraction

The masks of the aortic arch were manually segmented using Adobe Photoshop CC 2018 (Adobe Systems Inc., San Jose, CA, USA) and served as “ground truth”. The mask was created by an experienced radiologist, drawing along the boundary of the aortic arch in the raw images.

For automatic ROIC extraction. First, we found the center position of the target aortic arch using SegNet convolutional neural network as suggested by others. And then, the raw image was cropped so that the center of each aortic arch was centered to cover the entire aortic arch portion (Fig 5). Here, the size of the box was determined based on the largest area of each aortic arch in the mask image to be used for training (800 X 680 X 3 pixels).

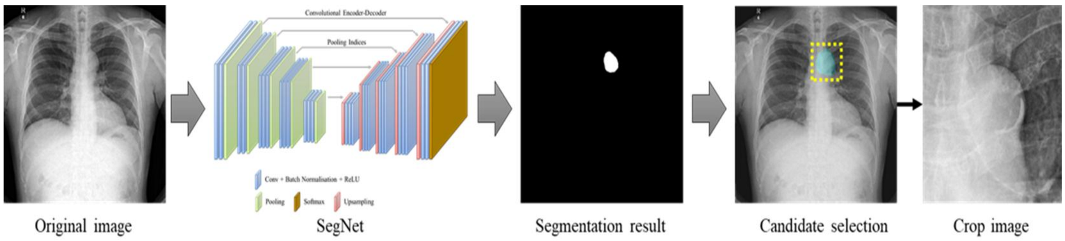


Figure 11. Method for automatic ROIC extraction

The SegNet model was trained with batch size of 4, maximum iteration number of 120 echoes, and learning rate of 0.01. The images were resized to 428 X 428 X 3 pixels with intensities scaled to [0,1].

The SegNet model was trained by using the training and validation data and implemented with MatLab R2018b on a GeForce GTX 1080Ti graphics processing unit.

C. Determination of aortic arch calcification grade

Alex-Net[23], ZF-Net[24], GoogLeNet[25], VGG-Net[26] and ResNet[27] has been used extensively as a pretrained network to classify images for the medical field.

Alex-Net is the first CNN to win the 2012 ImageNet Challenge. AlexNet's CNN consists of 5 Conv layers and 3 fully connected layers, with 96 to 384 filters and 3 to 256 channels in each Conv layer. ReLU nonlinearity is used at each layer. The maximum pooling of 3 × 3 applies to the outputs of layers 1, 2 and 5. Alex-Net used 4 strides in the first layer of the network.

For automatic determination of aortic arch calcification grade. First, we classified the aortic arch calcification grade into two major grades (grades 0 and 1 merged into one, grades 2 and 3 merged into one) using Alexnet convolutional neural network (Fig 6). The Alexnet model was trained with batch size of 1, maximum iteration number of 120 echoes, and learning rate of 5xe-4. The images were resized to 227 X 227 X 3 pixels with intensities scaled to [0,1].

The Alexnet model was trained using training and validation data and implemented in the same environment as the ROIC extraction algorithm development environment.

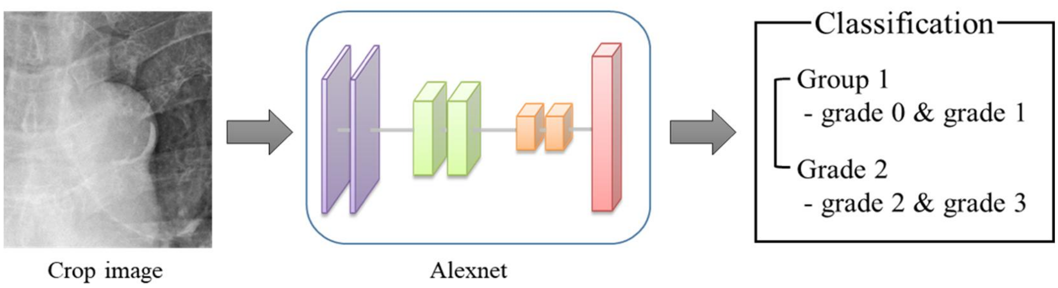


Figure 12. Determination of aortic arch calcification grade into two major grades. Next, the aortic arch calcification grade was subdivided by SVM(Support Vector Machine) based on HOG(Histogram of Oriented Gradients) feature (Fig 5).

The SVM is one of the fields of machine learning that creates a feature by learning the positive data and the negative data. And one of these features is the HOG Feature, which is a method of using the local gradient of the image as a feature of the image. In this study, for meaningful feature extraction, HOG features were extracted by defining a region for feature extraction in each group image, and then creating a binary image in which calcified regions are separated based on intensity values.

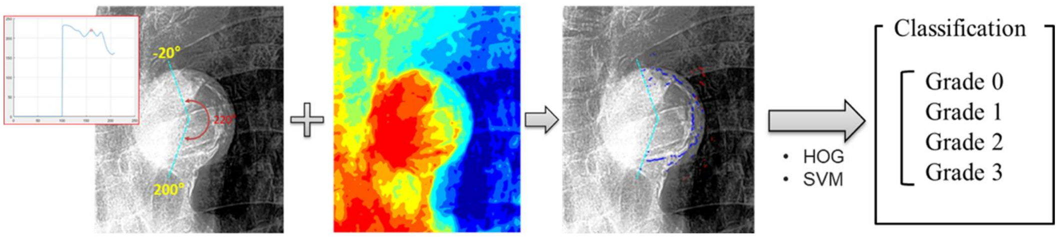


Figure 13. Subdivision of aortic arch calcification grade

5. Acquisition of carotid ultrasound and clinical information

Complete the database by writing the clinical findings from January 1, 2007 to March 31, 2012, common carotid artery IMT, internal carotid artery doppler results and clinical information in the form of a clinical report.

i. Carotid ultrasound results are checked in the medical record, and additional common carotid artery stiffness measurements are analyzed in the core lab of the reading center.

ii. Database parameter

1. Basic factors: height, weight, systolic phase of the patient at the time of carotid ultrasound diastolic blood pressure, pulse rate
2. Cardiovascular disease risk factors: high blood pressure, diabetes, smoking, hyperlipidemia, obesity (BSA, BMI), FRS score
3. Cardiovascular disease diagnosis: acute myocardial infarction, old myocardial infarction, unstable angina, stable angina, atypical chest pain, risk factor only Cardiovascular disease diagnosis:
4. Coronary arteriography findings: 1VD, 2VD, 3VD, Left main disease
5. Echocardiographic findings reflecting vascular stiffness: PWV, baPWV, ABI, CCA, TDIS, LVEF
6. Lab finding: total cholesterol, HDL cholesterol, LDL cholesterol, Triglyceride, Creatinine

7. Medication: aspirin, statin, angiotensin converting enzyme inhibitor, angiotensin receptor blocker, beta blocker, calcium channel blocker

iii. Outcome monitoring is collected by reviewing medical records up to September 2020 by researchers.

8. Primary end-points
   A. Cerebral infarction
   B. Cardiovascular accidents (all cardiovascular events; acute coronary syndrome, cardiovascular death)
   C. Cardio-cerebrovascular accidents (composite vascular events; cerebral infarction, acute coronary syndrome, cardiovascular death)
9. Secondary end-points
   A. Coronary artery reperfusion therapy (new coronary revascularization)
   B. Hospitalization due to heart failure (admission due to heart failure)
   C. The occurrence of atrial fibrillation (development of atrial fibrillation)

6. Statistical analysis

Data are demonstrated as the mean ± SD for continuous variable, and presented as number (N) and percent (%) for categorical variables. SAS version 9.4 (SAS Institute, Cary, NC, USA), R package version 4.0.3 (http://www.R-project.org) were used.

Demographics were analyzed with ANOVA according to AAC grade on CXR by naked eye for the intention to compare continuous variables, $\chi 2$ tests for categorical variables followed by post hoc analyses with the Bonferroni method. To evaluate the association between AAC grade on CXR and primary endpoints, we used Kaplan-Meyer curve with log rank test. The primary endpoints were CVA event, all cardiovascular event, treatment of PTCA, all cause death and hospitalization due to heart failure aggravation. We also analyzed the hazard ratio of each event according to aortic arch calcification groups with adjustment of FRS.

To evaluate predictability, Harrell's c-index was used. We compared three groups ; FRS alone , FRS with aortic arch calcification grade, aortic arch calcification grade alone.

Future statistical analysis would be as follows;

① Inter-observer reliability of AAC grade : weighted kappa

② Diagnostic ability of AAC grade between human and AI : Delong method, AUC area

III. RESULTS

At the preliminaries, total 3,080 patients were reviewed among 5,022 patients who had carotid Doppler ultrasound at Gangnam Severance Hospital from January 1, 2007 to March 31, 2012. Among them, 97 patients did not have chest x-ray. 710 patients were dropped because time difference between the date of chest x-ray and the date of carotid Doppler ultrasound was larger than 3 years. Hence, total 2,273 patients were analyzed at preliminaries.

The prevalence of each AAC grade on chest x-ray by naked eye is 45.5% for AAC grade 0, 28.6% for AAC grade 1, 22.0% for AAC grade 2 and 3.9% for AAC grade 3, respectively.

Individuals with older age, female, DM, any history of smoking, decreased renal function, higher FRS, history of heart failure and history of HTN had significantly association with higher aortic arch calcification grade (p value < 0.0001). Patient with MI history, AF and dyslipidemia was not associated with AAC grade on chest x-ray. Experience of any angina or stroke was statistically significant according to AAC grade on chest x-ray (p value 0.0187, 0.0007 respectively).

Table 3. Baseline characteristics of study population by AoAC grade on chest x ray

	AoAC Grade on CXR					*p*
	Total (n=3745)	Grade 0 (n=1779)	Grade 1 (n=998)	Grade 2 (n=820)	Grade 3 (n=148)	
Age	62.1±11.4	57.5±10.8	63.1±10.1	68.8±9.4	73.5±8.4	<0.0001§
Male (%)	2209 (59)	1136 (63.9)	597 (59.8)	424 (51.7)	52 (35.1)	<0.0001
Smoking						<0.0001
Current smoker		348 (54.5)	181 (18.7)	102 (12.5)	7 (4.8)	
Ex-smoker		415 (24.2)	239 (24.6)	191 (23.5)	27 (18.4)	
Never-smoker		952 (55.5)	550 (56.7)	520 (64.0)	113 (76.9)	
Systolic BP (mmHg)	127.3±19.3	125.6±18.4	127.4±19.2	129.9±20.1	133.7±21.6	<0.0001
Diastolic BP (mmHg)	76.1±12.3	76.7±12.3	76.6±12.4	74.6±12.2	73.2±11.1	<0.0001
Pulse Pressure (mmHg)	51.2±13.8	48.8±12.6	50.8±13.0	55.4±15.1	60.5±17.0	<0.0001
baPWV, Left	1574.6±372.6	1474.6±298.7	1600.6±391.6	1681.3±389.2	1918.1±478.7	<0.0001
baPWV, Right	1548.9±355	1457.7±285.7	1562.3±341.7	1661±403.9	1843.8±472.1	<0.0001
baPWV, average	1562.1±353.8	1466.3±285	1582.7±354.7	1671.6±383.8	1878.3±463.8	<0.0001
RVP (mmHg)	25.2±7.5	24.1±6.1	25.1±7.3	26.7±8.2	29.4±13.7	<0.0001
MDRD eGFR (ml/min*1.73m^2)	74.5±16.4	77.3±14.4	74.5±15.4	70.8±18.5	61.8±22.5	<0.0001
Total cholesterol (mg/dL)	166.2±38.7	168.8±39.2	165.2±39	163.9±37	154.7±38.3	<0.0001
HDL (mg/dL)	44.4±11.7	44.9±11.7	44.5±11.9	44±11.1	40.9±11.8	0.002
LDL (mg/dL)	100.1±33	101.5±33.6	99.2±33	99.4±31.6	92.2±32.5	0.018
FRS	23.8±18.5	19.9±16.4	24.9±18.2	29.4±20.3	31.4±22.4	<0.0001
Past History						
Hypertension	2499 (66.7)	1036 (58.2)	719 (72)	622 (75.9)	122 (82.4)	<0.0001
Diabetes	1032 (27.6)	388 (21.8)	280 (28.1)	291 (35.5)	73 (49.3)	<0.0001
MI	464 (12.4)	213 (12.0)	128 (12.8)	100 (12.2)	23 (15.5)	0.6
Dyslipidemia	758 (20.2)	347 (19.5)	206 (20.6)	174 (21.2)	31 (20.9)	0.751
Stroke	379 (10.1)	147 (8.3)	93 (9.3)	113 (13.8)	26 (17.6)	<0.0001
Heart failure	99 (2.6)	28 (1.6)	21 (2.1)	39 (4.8)	11 (7.4)	<0.0001§
Atrial fibrillation	178 (4.8)	77 (4.3)	49 (4.9)	42 (5.1)	10 (6.8)	0.51

§ *p* value by Chi-square test with Fisher's exact test

Abbreviations : AoAC grade is aortic arch calcification grade ; BP is Blood Pressure ; MDRD eGFR is estimated GFR by Modification of Diet in Renal Disease equation ; FRS is Framingham Risk Score; PWV is Pulse Wave Velocity ; ba is brachial ankle ; RVP is Right Ventricular Pressure ; MI is myocardial infarction;

Relationship between AAC on chest x-ray, carotid Doppler ultrasound findings and arterial stiffness index

To examine the role of AAC grade on chest x-ray, logistic regression analysis was used. The OR (odds ratio) with 95% CI of AAC grade of chest x-ray to affect the development of carotid plaque was 1.33 (1.21 < 95% CI < 1.47, p value <0.0001). Similar pattern was noted in terms of age. The OR of age in determination of carotid plaque was 1.08 (1.07 < 95% CI <1.09, p value <0.0001). R square was 0.212.

Result of crosstab analysis was in accordance with logistic regression analysis. Table 3 indicates strong relationship of AAC grade on chest x-ray and carotid plaque on carotid Doppler ultrasound. OR of carotid plaque in case of AAC grade 3, 2, 1 compared to AAC grade 0 is 5.6, 3.6 and 1.73, respectively (p value <0.0001).

Table 4. AoAC grade, carotid ultrasound findings and baPWV.

		AoAC Grade on CXR				P
		0 (n=1779)	1 (n=998)	2 (n=820)	3 (n=148)	
Any carotid plaque	Yes, n (%)	482 (35.1)	389 (48.3)	425(65.8)	87(75.0)	<0.001

Table 3. Relationship of AoAC and arterial stiffness through baPWV after adjustment for age

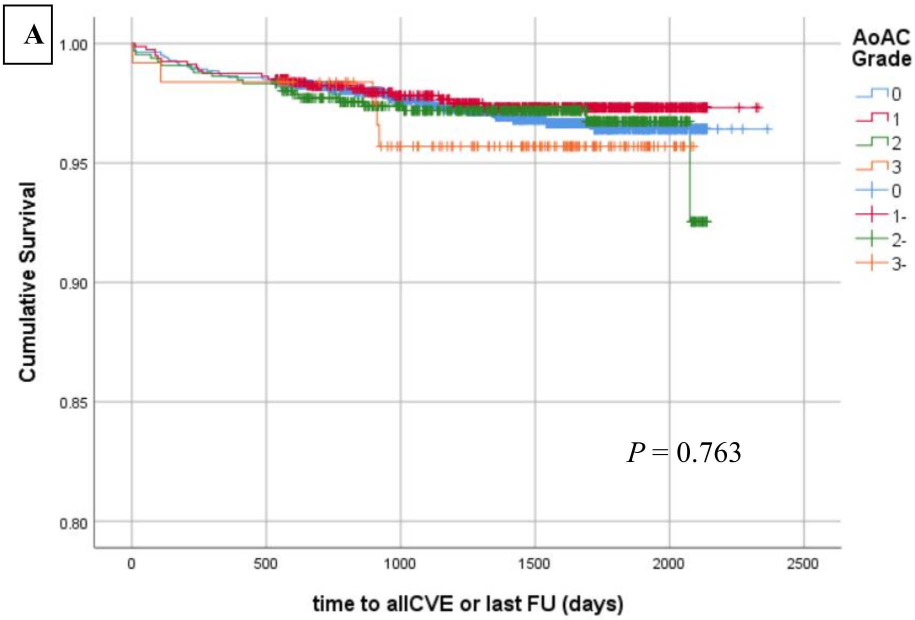
		Standardized Coefficients		95.0% Confidence Interval	
E/e'		**Beta**	Sig.	Lower Bound	Upper Bound
	(Constant)		0.000	1.512	3.376
	AAC_grade	0.134	0.000	0.483	0.865
	age	0.333	0.000	0.121	0.153

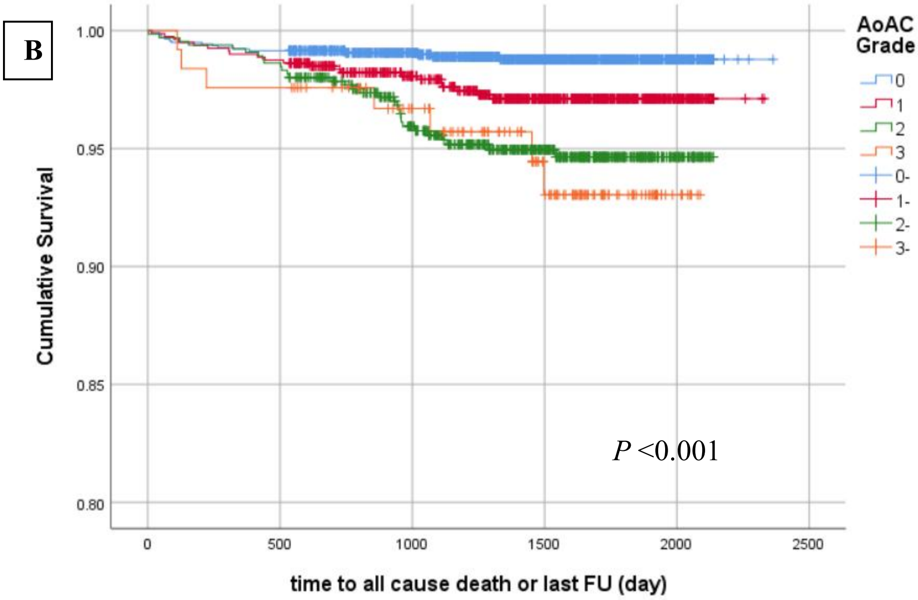
		Standardized Coefficients		95.0% Confidence Interval	
Rt. CCA-EDV		**Beta**	Sig.	Lower Bound	Upper Bound
	(Constant)		0.000	31.539	35.555
	AAC_grade	-0.034	0.115	-0.715	0.078
	age	-0.309	0.000	-0.274	-0.208

Representative marker of arterial stiffness such as baPWV and , TDIEE and EDV of Rt CCA were all significantly associated with AAC grade independent of age (Table 4).

Relationship between AAC grade on chest x-ray and cardio-cerebrovascular outcomes

Overall, all the outcomes, excepts for ischemic coronary events, were statistically related to AAC grade in Kaplan-Meier log rank test. Presence of AAC [graded as none and existence of calcification in any severity] were strongly associated with future CVA events and all cause death even after adjustment for FRS (Figure 7. A, B). However, treatment with PTCA and all CVE such as MI, unstable angina and all cause cardiovascular death were not associated with AAC grade (Figure 7 C, D). Hospitalization due to heart failure was only associated with AAC grade 3 (p = 0.003).


A
1.00
0.95
0.90
0.85
0.80
Cumulative Survival
0
500
1000
1500
2000
2500
time to allCVE or last FU (days)
AoAC Grade
0
1
2
3
0
1-
2-
3-
P = 0.763


B
1.00
0.95
0.90
0.85
0.80
Cumulative Survival
0
500
1000
1500
2000
2500
time to all cause death or last FU (day)
AoAC Grade
0
1
2
3
0-
1-
2-
3-
P <0.001

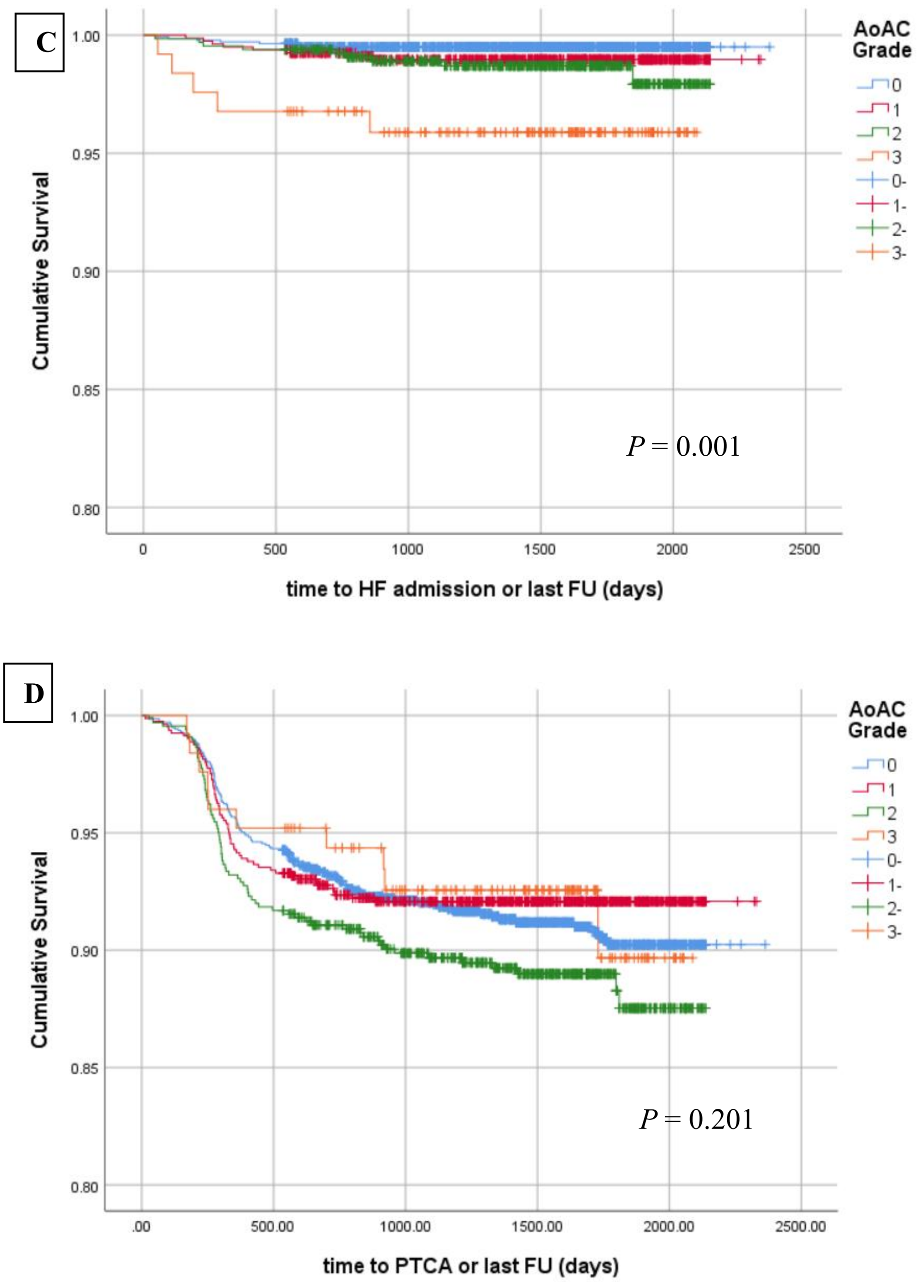


Figure 14. Kaplan Meier curve with log rank test for primary end-points

Table 4. Cox regression for cardio-cerebrovascular events after adjustment for HRS.

Variables	**All CVE**			
	Univariable model		Multivariable model	
	HR(95% CI)	p-value	HR(95% CI)	p-value
AAC grade				
0	ref		ref	
1	1.09(1.06-3.07)	0.03	1.63(0.90-3.00)	0.11
2	2.10(2.00-3.68)	0.009	1.74(0.93-3.25)	0.09
3	2.89(1.18-7.05)	0.02	2.00(0.68-5.91)	0.21
FRS			1.02(1.01-1.03)	<0.0001
	All_cause_death			
AAC grade				
0	ref		ref	
1	2.28(1.19-4.37)	0.013	2.32(1.04-5.18)	0.039
2	4.16(2.28-7.61)	<0.0001	4.40(2.10-9.20)	<0.0001
3	4.92(2.20-12.00)	<0.0001	5.39(1.93-15.00)	0.001
FRS			1.03(1.01-1.04)	<0.0001
	HF_admission			
AAC grade				
0	ref		ref	
1	1.593(0.559-4.542)	0.3836	2.72(0.79-9.31)	0.112
2	1.767(0.594-5.257)	0.3063	3.76(1.14-12.42)	0.03
3	8.393(2.664-26.444)	0.0003	9.53(2.35-38.68)	0.002
FRS			1.02(1.00-1.04)	0.008

Predictive power of AAC grade on chest x-ray; is it comparable with FRS ?

AAC grade alone showed significant predictive power for CVA event, all cause death and HF admission. We also compared FRS and AAC grade. C-index of AAC with FRS was higher than FRS alone, but it was not statistically significant (Table 6).

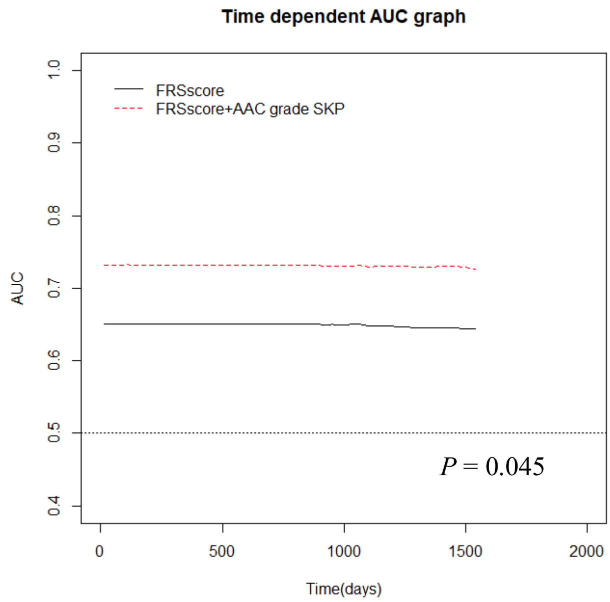


Figure 9. iAUC graph to compare the ability of AoAC grade and FRS to predict all-cause death

Deep learning algorithm for AAC grade

To evaluate the performance of segmentation algorithm, dice similarity coefficient (DSC) was implemented which uses the automated segmentation mask and the human-annotated segmentation mask.

DSC greater than or equal to 0.7 were considered as excellent agreement between two segmented regions in previous investigations, we defined the value of 0.7 or more as an appropriate segmentation performance.

As a result, the average of similarity was 62%. In this study, the segmentation algorithm is to extract the region of interest, not the exact shape segmentation. For this reason, we extracted the centroid positions of the automated segmentation mask and the human-annotated segmentation mask and calculated the distance error between the centroid positions. As a result, the average distance error is 3.42 pixels, which is considered a good result.

The classification performance of the aortic arch calcification grade was evaluated for each step.

First, the performance evaluation of the algorithm for classifying aortic arch calcification grades into two main grades are summarized in Table 1. At the operating point (threshold) of 0.5, sensitivity was 93.81% and 76.14% when specificity was 95.92% and 79.55% for the validation and test sets.

Next, the performance evaluation of the algorithm for subdividing aortic arch calcifications grades are summarized in Table 2. At the operating point (threshold) 0.5, the sensitivity was 72.7% and the specificity was 93.8% for the Group1 and the sensitivity was 86.4% and the specificity was 95.2% for the Group2.

Table 5. The performance evaluation of the algorithm for classifying aortic arch calcification grades into two main grades

	Sensitivity(%)	Specificity(%)	Accuracy(%)
Validation Set	93.81	76.14	85.41
Test Set	95.92	79.55	88.17

Table 6. the performance evaluation of the algorithm for subdividing aortic arch calcifications grades

	Sensitivity(%)	Specificity(%)	Accuracy(%)
Group1	72.7	93.8	83.1
Group2	86.4	95.2	90.7

IV. DISCUSSION

AAC as a surrogate marker for arteriosclerosis and aortic stiffness

The first purpose of this study was achieved. We confirmed that aortic arch calcification on chest x- ray was one of strong risk factors for CVA and all cause death. This result is consistent with previous studies [28, 29].

The implication of this study in evaluating aortic arch calcification grade on chest x ray is that the main mechanism of aortic arch calcification is arterial stiffness rather than atherosclerosis. We can witness the un-relevance statistical result between treatment of PTCA, all CVE and aortic arch calcification grade. Contrary to CVA, all cause death and HF admission, PTCA and all CVE were not associated with aortic arch calcification grade. We can conclude that coronary artery system, which was developed from myocardial cells are totally different from aortic arch, which was developed from angioblast. We can speculate that the difference of embryological difference can contribute the mechanism of vascular calcification. As myocardial cell derived coronary artery tend to be more tolerable to pressure than angioblast cell derived aortic arch.

Arterial stiffness is in line with PWV, IMT, and resistive index of carotid artery in Doppler ultrasound. We will analyze the Doppler ultrasound parameter after preliminaries.

Association of smoking and AAC is unclear. Most of studies demonstrated smoking was not significantly related to AAC on chest x-ray [29]. But other studies indicated smoking was relevant to AAC on chest x-ray [15] [30]. Iribarren et al. confirmed that smoking was independently associated aortic arch calcification on large, cohort study, composed of 116,319 people with median follow-up duration of 28 years. In our study, smoking history was not significantly related to AAC both in men and women. Interestingly, in our study smoking history was significantly related with history AMI, unstable angina, variant angina and PTCA but not with CVA, HF admission and atrial fibrillation (data not shown). In addition, the relationship between smoking and presence of carotid plaques was weak within each sex. It suggests that smoking affects more coronary atherosclerosis and plaque rupture than arteriosclerosis in aorta or peripheral arteries, which in line with our result of non-relationship between smoking history and AAC.

AAC as a prognostic factor or a component of risk prediction model

We recognized that AAC grade alone in chest x-ray is able to predict CVA, all-cause death and admission due to heart failure aggravation (Table 7). Compared with FRS alone, FRS with AAC grade, there was no significant advantage. However, the Harrell's c index increased when combined with FRS with AAC grade. This indicates that AAC grade has additive predictive value on cardio-cerebrovascular events. Statistically significance would be improved if we add more cases at the final report.

FRS was developed to predict coronary vascular event. Hence, it would be less reasonable to compare predictive power of FRS and AAC grade in all aspects of cardio-cerebrovascular outcome. According to our study, we explored that AAC grade on chest x-ray has predictive power to cardio-cerebrovascular outcome except coronary artery occlusive disease.

Feasibility of deep learning for AAC grading in chest x-ray

The current results are difficult to use in actual clinical practice, and for this purpose, it is necessary to improve the performance of the developed algorithm. In future research, we plan to improve the performance of the developed algorithm by studying a method that can add high-quality data and define more clear features.

We demonstrated the way of ROIC segmentation using deep learning algorithm. Although the algorithm still confused with AAC grade 0 and 3, we expect that the problem would be solved. We will use more training individuals who have non contrast chest CT, so that we can exactly code the true AAC grade 0. In this way, the algorithm can explicitly recognize grade 0. Similarly, algorithm can easily differentiate AAC grade 3 owing to its unique circular shape. So we will train the algorithm to divide the ROIC into AAC grade 0 and 1 VS AAC grade 2 and 3. After this first step, grade 1 is remnant of the small group except AAC grade 0, which is trained exactly through the true value using non contrast matched value. The other small group, containing AAC group 2, 3 can be differentiated after selecting AAC grade 3.

Future perspectives of deep learning in population based CV risk prediction model

Aortic arch calcification on chest x ray has its implication in that it tells us arterial stiffness which eventually leads to difficult volume control in hospitalized patient. Fortunately, we concluded that the arterial stiffness through deep learning algorithm without complex equations such as FRS. Most of hospitalized patients usually undergo chest x ray at least when they need being hospitalized. Physicians can intuitively catch the risk of CVA, all cause death and HF aggravation by just checking the AAC grade on x ray. Physician have another opportunity to order fluid with delicate attention and discuss about the patient with internist efficiently if a physician has insufficient idea on fluid treatment.

Cardio-cerebrovascular accidents usually occur with interaction with each other. And it is hard to observe that the cardio-cerebrovascular accident happens just one time. The tendency of recurrence of cardio-cerebrovascular accident emphasize the implication of this study. Obtaining the insights of cardio-cerebrovacsular accidents with simple tool such as x ray will provide the opportunities for the physicians to attainment of early intervention of internists. This is the first study that statistically evaluated the relationship between aortic arch calcification, carotid Doppler ultrasound parameters and echocardiographic values.

FRS has been a useful and relatively accurate tool for prediction of 10 year coronary artery disease, but has limited implication because FRS does not deal with broad cardio-cerebrovascular problems and need to draw blood sample to calculate FRS. Machine learning (ML) is area of computer data science that manipulate algorithms to recognize patterns in huge datasets with many variables. Through this process, ML can expect and predict outcomes that researchers are interested. Till now, the advantage of ML from CXR AAC to predict cardio-cerebrovascular events has not been investigated yet. Therefore this study to search for the feasibility of ML algorithm to grade AAC on CXR to predict cardio-cerbrovascular events guarantees future related studies.

REFERENCES

1. National Statistical Office, *Analysis of cause of death in 2018*, Vital Statistics Division, Editor. 2019: Sejong.
2. Jiaquan Xu, M.D., Sherry L. Murphy, B.S., Kenneth D. Kochanek, M.A., and Elizabeth Arias, Ph.D., *Mortality in the United States, 2018*, Division of Vital Statistics, Editor. 2020, National Center for Health Statistics: Hyattsville, MD.
3. Centers for Disease Control, *Death and Mortality* in *FactStats Homepage* National Center for Health Statistics, Editor. 2017.
4. *Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report.* Circulation, 2002. **106**(25): p. 3143-421.
5. D'Agostino, R.B., et al., *General Cardiovascular Risk Profile for Use in Primary Care.* Circulation, 2008. **117**(6): p. 743-753.
6. Goff, D.C., et al., *2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines.* Journal of the American College of Cardiology, 2014. **63**(25, Part B): p. 2935-2959.
7. Goff, D.C., et al., *2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk.* Circulation, 2014. **129**(25_suppl_2): p. S49-S73.
8. ClinCalc. *ASCVD Risk Calculator* Cardiology 2013 Nov 2013 [cited 2020 Oct 4];

Available from: https://clincalc.com/Cardiology/ASCVD/PooledCohort.aspx.

9. *Joint British Societies' consensus recommendations for the prevention of cardiovascular disease (JBS3).* Heart, 2014. **100**(Suppl 2): p. ii1.
10. Flueckiger, P., et al., *Revised Framingham Stroke Risk Score, Nontraditional Risk Markers, and Incident Stroke in a Multiethnic Cohort.* Stroke, 2018. **49**(2): p. 363-369.
11. Golestani, R., et al., *Abdominal aortic calcification detected by dual X-ray absorptiometry: A strong predictor for cardiovascular events.* Ann Med, 2010. **42**(7): p. 539-45.
12. Jayalath, R.W., S.H. Mangan, and J. Golledge, *Aortic calcification.* Eur J Vasc Endovasc Surg, 2005. **30**(5): p. 476-88.
13. Ma, X., et al., *Aortic Arch Calcification Is a Strong Predictor of the Severity of Coronary Artery Disease in Patients with Acute Coronary Syndrome.* BioMed Research International, 2019. **2019**: p. 7659239.
14. Ma, X., et al., *Predictive performance of aortic arch calcification for clinical outcomes in patients with acute coronary syndrome that undergo percutaneous coronary intervention: A prospective clinical study.* Medicine (Baltimore), 2019. **98**(48): p. e18187.
15. Hashimoto, H., et al., *Validity and usefulness of aortic arch calcification in chest X-ray.* J Atheroscler Thromb, 2009. **16**(3): p. 256-64.
16. Iijima, K., et al., *Aortic arch calcification detectable on chest X-ray is a strong independent predictor of cardiovascular events beyond traditional risk factors.*

Atherosclerosis, 2010. **210**(1): p. 137-44.

17. Pezzotti, W., *Chest X-ray interpretation: not just black and white.* Nursing, 2014. **44**(1): p. 40-7; quiz 47-8.
18. Woznitza, N., et al., *Chest X-ray Interpretation by Radiographers Is Not Inferior to Radiologists: A Multireader, Multicase Comparison Using JAFROC (Jack-knife Alternative Free-response Receiver Operating Characteristics) Analysis.* Acad Radiol, 2018. **25**(12): p. 1556-1563.
19. Spaedy, E., et al., *Accuracy of remote chest X-ray interpretation using Google Glass technology.* International Journal of Cardiology, 2016. **219**: p. 38-40.
20. Medicine, N.L.o. *Search Query result by "chest x ray machine learning"*. 2020 [cited 2020 Sep 14]; Available from: https://pubmed.ncbi.nlm.nih.gov/?holding=ikryumlib&otool=ikryumlib&term=chest+x+ray+machine+learning.
21. Filice, R.W., et al., *Crowdsourcing pneumothorax annotations using machine learning annotations on the NIH chest X-ray dataset.* J Digit Imaging, 2020. **33**(2): p. 490-496.
22. Badrinarayanan V, K.A., Cipolla R. SegNet, *A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Trans Pattern.* Anal Mach Intell 2017. **39**: p. 2481-2495.
23. Krizhevsky A, S.I., Hinton GE, *Advances in neural information processing systems*. Imagenet classification with deep convolutional neural networks. 2012.
24. Zeiler MD, F.R., *European conference on computer vision Springer*. Visualizing and

understanding convolutional networks. 2014.

25. Szegedy C, L.W., Jia Y, Sermanet P, Reed S, Anguelov D, et al, *Proceedings of the IEEE conference on computer vision and pattern recognition*. Going deeper with convolutions. 2015.

26. Simonyan K, Z.A., *Very deep convolutional networks for large-scale image recognition*, in *ICLR 2015*. 2014, arXiv.

27. He K, Z.X., Ren S, Sun J, *Proceedings of the IEEE conference on computer vision and pattern recognition*. Deep residual learning for image recognition. 2016.

28. Hollander, M., et al., *Comparison between measures of atherosclerosis and risk of stroke: the Rotterdam Study.* Stroke, 2003. **34**(10): p. 2367-72.

29. Hendriks, E.J.E., et al., *Annularity of Aorto-Iliac Arterial Calcification and Risk of All-Cause and Cardiovascular Mortality.* JACC Cardiovasc Imaging, 2018. **11**(11): p. 1718-1719.

30. Iribarren, C., et al., *Calcification of the aortic arch: risk factors and association with coronary heart disease, stroke, and peripheral vascular disease.* Jama, 2000. **283**(21): p. 2810-5.