

Interpretation of captopril renography using artificial neural networks

Magnus Nielsen¹, Göran Granerus², Mattias Ohlsson^{1,3}, Holger Holst¹, Ola Thorsson¹ and Lars Edenbrandt^{1,4}

¹Department of Clinical Sciences, Malmö University Hospital, Malmö, ²Department of Clinical Physiology, Linköping University Hospital, Linköping,

³Department of Theoretical Physics, Lund University, Lund, and ⁴Department of Clinical Physiology, Sahlgrenska University Hospital, Göteborg, Sweden

Summary

Correspondence

Lars Edenbrandt, Department of Clinical Sciences, Malmö, Lund University Research Program in Medical Informatics, LUMI, UMAS, Entrance 44, SE-205 02 Malmö, Sweden
E-mail: lars.edenbrandt@med.lu.se

Accepted for publication

Received 4 April 2005;
accepted 10 June 2005

Key words

captopril renography; computer-assisted diagnosis; hypertension; radioisotope renography; radionuclide imaging; renal artery obstruction; renovascular

The purpose of this study was to develop a method based on artificial neural networks for interpretation of captopril renography tests for the detection of renovascular hypertension caused by renal artery stenosis and to assess the value of different measurements from the test. A total of 250 99mTc-MAG3 captopril renography tests were used in the study. The material was collected from two different patient groups. One group consisted of 101 patients who also had undergone a renal angiography. The angiographies, which were used as gold standard, showed a significant renal artery stenosis in 53 of the 101 cases. The second group consisted of 149 patients, who's captopril renography tests all were interpreted as not compatible with significant renal artery stenosis by an experienced nuclear medicine physician. Artificial neural networks were trained for the diagnosis of renal artery stenosis using eight measures from each renogram. The neural network was then evaluated in separate test groups using an eightfold cross validation procedure. The performance of the neural networks, measured as the area under the receiver operating characteristic curve, was 0.93. The sensitivity was 91% at a specificity of 90%. The lowest performance was found for the network trained without use of a parenchymal transit measure, indicating the importance of this feature. Artificial neural networks can be trained to interpret captopril renography tests for detection of renovascular hypertension caused by renal artery stenosis. The result almost equals that of human experts shown in previous studies.

Introduction

Renal angiography is the gold standard for diagnosing renovascular disease. It is an invasive method and not suitable for screening. The most accurate non-invasive method to diagnose significant renal artery stenosis causing renovascular hypertension (RVH) is the captopril renography test (CRT) (Elliott *et al.*, 1993; Fommei *et al.*, 1993; Pedersen *et al.*, 1996; Helin *et al.*, 1998), which is a well-established technique since almost two decades. A consensus report (Taylor *et al.*, 1996) and procedure guidelines (Taylor *et al.*, 1998) have been presented to assist nuclear medicine physicians in performing and interpreting the outcome of the test. These guidelines for interpretation of the CRT are based on the possibility to compare the captopril study to a baseline study in case of an abnormal CRT. No consensus, however, has yet been reached for interpretation of the individual output features from the CRT. It is often possible to perform only the captopril study and there are no definite

guidelines to determine when a baseline study is necessary. According to one report (Ramsay *et al.*, 1997) clinicians sometimes rely on other factors than the CRT when determining who has a possible symptomatic stenosis and may benefit from intervention, despite the fact that CRT is both sensitive and specific for RVH. It is also possible that some clinics use their own diagnostic criteria that are not statistically evaluated. Krijnen *et al.* (2002) investigated the interobserver agreement on CRT and they concluded that differences in the performance of CRT found between studies could be explained by this variability. The assessment of cortical retention and pelvic retention by visual inspection of the images showed the largest variability.

Artificial neural networks have been used for the purpose of diagnosis in nuclear medicine, for example ischaemic heart disease, pulmonary embolism and Alzheimer's disease (Page *et al.*, 1996; Lindahl *et al.*, 1999; Holst *et al.*, 2001). The networks have proved to be well suited to solve pattern

recognition tasks and it has been shown that even an experienced expert benefit from the advice of artificial neural networks (Lindahl et al., 1999).

The purpose of this study was to develop a method based on artificial neural networks for interpretation of CRTs for the detection of RVH caused by renal artery stenosis and to assess the value of different measurements from the test.

Methods

Patient selection

This study is based on technetium-99m-mercaptoacetyltri-glycine (99m-Tc-MAG3) CRTs performed on patients suspected to suffer from RVH. The material was collected from two different patient groups. One was part of the Swedish Society of Nuclear Medicine's Captopril Renography Project. It comprised 105 tests performed 1995–1999 at a number of hospitals in Sweden according to guidelines proposed by the Swedish Society and very similar to the procedure guideline published by Society of Nuclear Medicine USA (Taylor et al., 1998). Each patient in this group has also undergone a renal angiography. Three tests were excluded because the evaluations of the renographies were not complete and one because of a non-conclusive angiography report. The criteria used for a positive angiography was a narrowing of the lumen diameter of more than 70% or signs of fibro-muscular dysplasia on one or both of the renal arteries. A positive angiography was found in 53 of the 101 tests (Fig. 1).

The second group comprised all patients who performed a CRT during a 12-month period 1998 at Linköping University Hospital. One experienced nuclear medicine physician interpreted all tests and included in our material was the 149 cases where he ruled out significant renal artery stenosis. There was a wide range of renal function in this group of patients, ranging from 20 to 570 ml min⁻¹ 1.73 m⁻² BSA as calculated as camera based MAG3 clearance (Granerus & Moonen, 1991). These patients did not undergo renal angiography and the interpreta-

tion of the expert was used as the gold standard. At a follow-up 5 years after the CRTs none of the 149 patients had undergone renal angiography.

The total material comprised 250 CRTs, 53 positive and 197 negative for renal artery stenosis. All of the patients had two native kidneys at the time.

The captopril renography test

Although the CRTs were performed at different hospitals a common protocol was recommended and the following minimal requirements were agreed on. The patients should be hydrated with 7 ml kg⁻¹ of body weight and 25 mg of captopril administered orally 60 min prior to the study. 99m-Tc-MAG3, 50–100 MBq, should be injected with the patient lying in the supine position on the gamma camera detector. The acquisition ought to be 20 min and 10-s frames acquired at least in the first part of the study. Experienced operators should process the images. The following measurements from the CRT were used:

- (i) relative function;
- (ii) time to highest activity (T_{max});
- (iii) activity after 20 min to highest activity (min 20/max);
- (iv) Pelvic appearance time (PAT).

A general purpose collimator was used and acquisition was made in 128 × 128 matrix. PAT was assessed by one experienced nuclear medicine physician who determined in which 1-min image the isotope first was visible in the renal pelvis. Four measurements were used for each of the right and left kidney, i.e. in total eight measurements. The artificial neural networks does not make use of normal limits for the measurements and such limits were therefore not defined.

Artificial neural networks

Artificial neural networks were used as classifiers for the detection of RVH. The analysis was done per patient, i.e. the classifiers computed the probability of RVH for each patient, thereby including left, right and bilateral stenosis. Each classifier consisted of an ensemble of single artificial neural networks. The individual members of the ensemble were standard multi-layer perceptrons (Rumelhart & McClelland, 1986) with one hidden layer consisting of four nodes and one output node that was used to encode the presence of RVH or not. Each multi-layer perceptron was trained using gradient descent applied to a cross-entropy error function. The gradient descent method was augmented with a traditional momentum term and a Langevin extension (Rögnvaldson, 1994). To avoid over-training a weight elimination (Hanson & Pratt, 1989) regularization term was utilized. The output of the neural network ensemble was computed as the mean of the output of the individual members in the ensemble. In this study, an ensemble size of 100 multi-layer perceptrons was used.

In order to assess the relative importance of the different measurements from each CRT, four networks were trained and

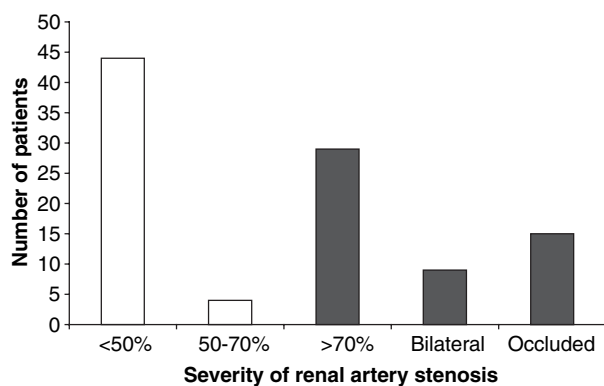


Figure 1 Severity of renal artery stenosis in the 101 patients who had undergone renal angiography. Black bars indicate the group of 53 patients with significant stenosis.

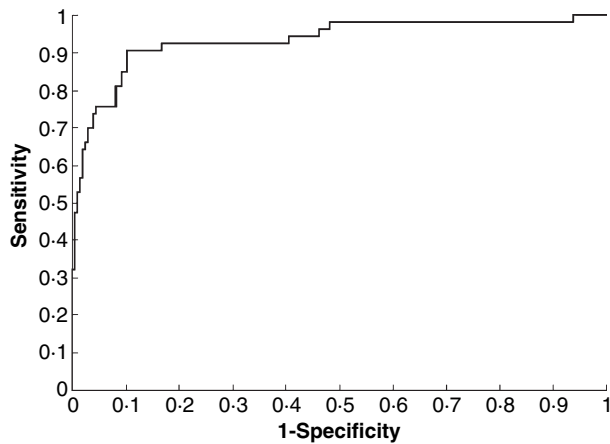


Figure 2 ROC curve presenting the performance of the neural networks interpreting CRTs based on eight measurements from each test.

evaluated using the same material but different inputs. For each of these networks, one of the four types of CRT measurements were omitted, for example the T_{\max} measurements for the right and left kidney. The performance of the neural network classifiers was measured as the area under the receiver operating characteristic (ROC) curve. All figures are based on eightfold cross-validation which was used as the technique to estimate the generalization performance. To find statistical significant differences between ROC areas, *P*-values were computed using a bootstrap method (Wehrens *et al.*, 2000).

Results

The performance of the neural network fed with all measurements from each CRT had an area under the ROC curve of 0.93 (Fig. 2). The sensitivity was 91% at a specificity of 90%. The networks trained without relative function, T_{\max} , and min 20/max showed areas under the ROC curves of 0.94, 0.94, and 0.92, respectively. The lowest area under the ROC curve, 0.89, was found for the network trained without the PAT measurements. This network was significantly worse than the network trained without T_{\max} measurements ($P = 0.02$). The other differences between the networks were not statistically significant.

Discussion

Main findings

This study shows that it is possible to use an artificial neural network to interpret CRTs in the detection of RVH. The accuracy is about as high as for human experts shown in other studies (Gijsbert & deBruyn, 1991; Mann *et al.*, 1991; Setaro *et al.*, 1991; Roccatello *et al.*, 1992; Fommei *et al.*, 1993), i.e. sensitivities and specificities in the region of 90%, respectively. It has been shown before by Hamilton *et al.*, 1996, that it could be done for transplant kidneys. This is to our knowledge the first time it has been shown to diagnose RVH in patients with native

kidneys. Furthermore, it is shown that the performance is high without the use of a baseline study.

One of the goals was to assess the value of different measurements from the CRT. The simple PAT showed to be the most important measurement and this finding indicates the importance of the parenchymal transit measurements. This feature was manually determined by an experienced nuclear medicine physician, however, it could be more objectively determined by deconvolution of cortical renograms and calculation of the parenchymal transit time (PTT). The three other types of measurements were standard features determined from the renograms. This suggests the mandatory use of some measurements of parenchymal transit could improve the performance of CRT, whether interpreted by computers or human experts. From a methodological point of view it is a great advantage for this purpose to use a tubular secreted tracer like MAG3 instead of a filtered tracer such as DTPA because of more rapid excretion and higher concentration in the pelvic urine. The higher counting statistics makes parameters like PAT or PTT more easily determined.

Study limitation

The performance of artificial neural networks depends heavily on the size and composition of the training database. The ideal situation would be to have a large number of CRTs from patients who all have undergone renal angiography. Because of the already high accuracy of CRT, renal angiography is a seldom-used investigation, especially for patients with a negative outcome of the CRT. Therefore, our second patient group used CRT interpreted by a human expert as gold standard rather than angiography. We used this database to increase the number of training examples. The CRT has, however, in many studies proved to be highly specific for a negative test (Gijsbert & deBruyn, 1991; Mann *et al.*, 1991; Setaro *et al.*, 1991; Roccatello *et al.*, 1992; Fommei *et al.*, 1993) and indeed the negative 5-year follow-up indicate that the expert interpretations were correct.

An more accurate gold standard than renal angiography would be the outcome of intervention to determine if a functional stenosis is present. This will give us the possibility to predict if a patient would gain from angioplasty. That information, however, was not available to us, but many captopril studies show that the CRT is a good predictor of the outcome (Gijsbert & deBruyn, 1991; Setaro *et al.*, 1991; Harward *et al.*, 1995).

Conclusion

This study shows that it is possible to use artificial neural networks for interpretation of CRTs in a wide range of renal function. The results are similar to those of human experts shown in earlier studies (Gijsbert & deBruyn, 1991; Mann *et al.*, 1991; Setaro *et al.*, 1991; Roccatello *et al.*, 1992; Fommei *et al.*, 1993), i.e. sensitivity and specificity in the region of 90%, respectively.

Acknowledgments

This study was supported by grants from the Swedish Medical Research Council (09893).

References

- Elliott WJ, Martin WB, Murphy MB. Comparison of two noninvasive screening tests for renovascular hypertension. *Arch Intern Med* (1993); **153**: 755–764.
- Fommei E, Ghione S, Hilson AJ, Mezzasalma L, Oei HY, Piepsz A, Volterrani D. Captopril radionuclide test in renovascular hypertension: a European multicentre study. European Multicentre Study Group. *Eur J Nucl Med* (1993); **20**: 617–623.
- Gijsbert GG, deBruyn AJG. Captopril Renography and the effect of percutaneous transluminal angioplasty on blood pressure in 94 patients with renal artery stenosis. *Am J Hypertens* (1991); **4**: 685S–689S.
- Granerus G, Moonen M. Effects of extra-renal background subtraction and kidney depth correction in the measurement of GFR by gamma camera renography. *Nucl Med Commun* (1991); **12**: 519–527.
- Hamilton D, Miola UJ, Mousa D. Interpretation of captopril transplant renography using a feed forward neural network. *J Nucl Med* (1996); **37**: 1649–1652.
- Hanson SJ, Pratt LY. Comparing biases for minimal network construction with backpropagation. In: *Advances in neural information processing systems* (ed. Touretzky, DS) (1989), pp. 177–185. Morgan Kaufmann, San Mateo, CA.
- Harward TR, Poindexter B, Huber TS, Carlton LM, Flynn TC, Seeger JM. Selection of patients for renal artery repair using captopril testing. *Am J Surg* (1995); **170**: 183–187.
- Helin KH, Tikkanen I, von Knorring JE, Lepantalo MJ, Liewendahl BK, Laasonen LS, Fyhrquist FY, Tikkanen T. Screening for renovascular hypertension in a population with relatively low prevalence. *J Hypertens* (1998); **16**: 1523–1529.
- Holst H, Mare K, Jarund A, Astrom K, Evander E, Tagil K, Ohlsson M, Edenbrandt L. An independent evaluation of a new method for automated interpretation of lung scintigrams using artificial neural networks. *Eur J Nucl Med* (2001); **28**:1, 33–38.
- Krijnen P, Oei HY, Claessens RA, Roos JC, van Jaarsveld BC, Habbema JD. Interobserver agreement on captopril renography for assessing renal vascular disease. *J Nucl Med* (2002); **43**: 330–337.
- Lindahl D, Lanke J, Lundin A, Palmer J, Edenbrandt L. Improved classifications of myocardial bull's-eye scintigrams with computer-based decision support system. *J Nucl Med* (1999); **40**: 96–101.
- Mann SJ, Pickering TG, Sos TA, Uzzo RG, Sarkar S, Friend K, Rackson ME, Laragh JH. Captopril renography in the diagnosis of renal artery stenosis: accuracy and limitations. *Am J Med* (1991); **90**: 30–40.
- Page MP, Howard RJ, O'Brien JT, Buxton-Thomas MS, Pickering AD. Use of neural networks in brain SPECT to diagnose Alzheimer's disease. *J Nucl Med* (1996); **37**: 195–200.
- Pedersen EB, Egeblad M, Jorgensen J, Nielsen SS, Spencer ES, Rehling M. Diagnosing renal artery stenosis: a comparison between conventional renography, captopril renography and ultrasound Doppler in a large consecutive series of patients with arterial hypertension. *Blood Press* (1996); **5**: 342–348.
- Ramsay D, Belton I, Finlay D. A review of captopril renal scintigraphy and its effect on patient management. *Nucl Med Commun* (1997); **18**: 631–633.
- Roccatello D, Picciotto G, Rabbia C, Pozzato M, De Filippi PG, Piccoli G. Prospective study on captopril renography in hypertensive patients. *Am J Nephrol* (1992); **12**: 406–411.
- Rögnvaldson T. On Langevin updating in multilayer perceptrons. *Neural Comput* (1994); **6**: 916–926.
- Rumelhart DE, McClelland JL (eds). *Parallel Distributed Processing*, Vol. 1 and 2 (1986). MIT Press, Cambridge, MA.
- Setaro JF, Chen CC, Hoffer PB, Black HR. Captopril renography in the diagnosis of renal artery stenosis and the prediction of improvement with revascularization. *Am J Hypertens* 1991; **4**: 698S–705S.
- Taylor A, Nally J, Aurell M, Blaufox D, Dondi M, Dubovsky E, Fine E, Fommei E, Geyskes G, Granerus G, Kahn D, Morton K, Oei HY, Russell C, Sfakianakis G, Fletcher J. Consensus report on ACE inhibitor renography for detecting renovascular hypertension. *J Nucl Med* (1996); **37**: 1876–1882.
- Taylor AT Jr, Fletcher JW, Nally JV Jr, Blaufox MD, Dubovsky EV, Fine EJ, Kahn D, Morton KA, Russell CD, Sfakianakis GN, Aurell M, Dondi M, Fommei E, Geyskes G, Granerus G, Oei HY. Procedure guideline for diagnosis of renovascular hypertension. *J Nucl Med* (1998); **9**: 1297–1302.
- Wehrens R, Putter H, Buydens LMC. The bootstrap: a tutorial. *Chemometr Intell Lab Syst* (2000); **54**: 35–52.