



Evaluation of Various Free Software Options for Catphan 504 Phantom Analysis

Fernandes^{a,b,*}, L.C.; Santos^a, M. R.; Silva^a, L. P.; Lima^{c,d}, T. V. M.; Simões^a, R. F. P.

^a Instituto Nacional de Câncer, Cruz Vermelha square, 23 - Centro, Rio de Janeiro -RJ, Brazil, 20230-130

^b Departamento de Engenharia Nuclear, Universidade Federal de Minas Gerais, Antônio Carlos Avenue, 6627 - Pampulha - Belo Horizonte, Brazil, 31270-901

^c Radiology and Nuclear Medicine, Luzerner Kantonsspital, Lucern, Switzerland

^d Department of Health Science and Medicine, University of Lucerne, Lucern, Switzerland

*Correspondence: lorena.fernandes@edu.inca.gov.br

Abstract: In computed tomography, image quality tests are important to guarantee a correct medical diagnosis and a better cost and benefit for the patient. Purpose: the purpose of this study is to analyse the images reconstructed with different thorax and bone convolution filters using popular free-use software in the field of medical physics, for the Catphan 504 phantom. Methods: a total of 14 scans were performed using the chest protocol, with convolution filters (FC) FC30, FC35, FC50, FC51, FC52, FC53, FC55, FC56, FC80, FC81, FC83, FC84, FC85, FC86, on the 16-channel Canon Aquilion Lightning CT scanner using the Catphan 504 phantom. Image quality parameters evaluated were: noise, uniformity, linearity of CT numbers, and high spatial resolution with MTF 50% and MTF 10%. The images were evaluated using software such as ImageJ, Script Python, and free software for automatic evaluation of the Catphan 504 phantom, CTQA-cp, SPICE-CT, and Pylinac. Results: the tests carried out with the Catphan 504 phantom were analysed by the software and agreed with each other (with $p > 0.05$), except for Pylinac. The results obtained with Pylinac had a significant difference for the uniformity, slice thickness, and MTF10% tests, this being the code that was the furthest away from the results obtained by the other codes. Conclusions: the ImageJ, Spice-CT, and CTQA-cp software showed consistent results for the tests performed, while Pylinac had limitations in calculating the standard deviation for the noise test and showed significant differences in some tests when compared to the other software.

Keywords: Computed Tomography, Catphan 504 phantom, Image quality, Free software analysis.



Avaliação de Diferentes Softwares Livres para Análise do Objeto Simulador Catphan 504

Resumo: Na tomografia computadorizada (TC), os testes de qualidade de imagem são importantes para garantir um diagnóstico médico correto e um melhor custo-benefício para o paciente. **Objetivo:** o objetivo deste estudo é analisar as imagens reconstruídas com diferentes filtros de convolução de tórax e ossos usando softwares livres no campo da física médica, para o objeto simulador Catphan 504. **Métodos:** um total de 14 varreduras foram realizados usando o protocolo de tórax, com filtros de convolução (FC) FC30, FC35, FC50, FC51, FC52, FC53, FC55, FC56, FC80, FC81, FC83, FC84, FC85, FC86, no scanner CT Canon Aquilion Lightning de 16 canais usando o objeto simulador Catphan 504. Os parâmetros de qualidade de imagem avaliados foram: ruído, uniformidade, espessura de corte, linearidade dos números de TC e alta resolução espacial com MTF 50% e MTF 10%. As imagens foram avaliadas usando os softwares livres, ImageJ, Script Python, CTQA-cp, SPICE-CT e Pylinac, para avaliação automática do objeto simulador Catphan 504. **Resultados:** os testes realizados com o Catphan 504 foram analisados pelos softwares e concordaram entre si (com $p > 0,05$), exceto para o Pylinac. Os resultados obtidos com o Pylinac tiveram uma diferença significativa nos testes de uniformidade, espessura de corte e MTF10%. **Conclusões:** os softwares ImageJ, Spice-CT e CTQA-cp mostraram resultados consistentes para os testes realizados, enquanto o Pylinac teve limitações no cálculo do desvio padrão para o teste de ruído e apresentou diferenças significativas em alguns testes quando comparado aos outros softwares.

Palavras-chave: Tomografia Computadorizada, Catphan 504, Qualidade de imagem, Softwares livres.

1. INTRODUCTION

The use of computed tomography (CT) imaging has been on the rise for several years, making quality assurance of CT exams a crucial requirement. This is in line with international recommendations and regulations in Brazil, which mandate that these exams must be performed and evaluated by trained and legally qualified professionals. [1-2].

When high-resolution imaging is necessary for accurate diagnosis, adjusting parameters in data processing can be an effective solution. Data processing is done using convolution filters (FCs), which can reduce artefacts and noise without increasing the radiation dose for the patient [3].

The assessment of image quality in CT scans involves the use of appropriate phantoms, such as the Catphan 504 (The Phantom Laboratory, Salem, NY, USA) or those provided by the CT manufacturer. These phantoms can include structures for measuring various image quality parameters such as uniformity, noise, slice thickness, low-contrast and high-contrast resolution, CT number linearity for different materials, and others. When these parameters do not meet the limits recommended by regulations or literature, it may indicate poor image quality [1, 2, 4].

To evaluate the images, there are several free software available for processing the Catphan 504 phantom [5, 6, 7, 8], such as ImageJ, Pylinac, Spice-CT, and CTQA-cp. These software can help standardize the analysis, minimize errors, and increase efficiency for the user. However, it remains to be determined if the results obtained by these software are consistent with each other and if they match the manual analysis of the phantom.

The objective of this study is to investigate the use of free software (ImageJ, Pylinac, Spice-CT, CTQA-cp) for analysing the Catphan 504 phantom, and to assess their agreement with each other in different convolution filters for thorax and bone imaging.

2. MATERIALS AND METHODS

In this study, scans were conducted using the Aquilion Lightning Canon CT scanner with 16 channels and the Catphan 504 phantom (The Phantom Laboratory, Salem, NY, USA), as illustrated in Figure 1. The images were analysed in DICOM format using free software: Spice-CT (V.0.3.4), Pylinac (V.3.1.0), CTQA-cp, ImageJ (V.1.8.0) and Python (V.3.6) [9, 10, 11, 12].

Figure 1: A) Catphan 504 phantom positioned on the Aquilion Lightning Canon TC scanner; B) and C) Catphan 504 phantom.



In total, 14 routine chest scans were performed, and for each scan, the convolution filter (FC) was changed. The protocol parameters were fixed: helical protocol, slice thickness 2.0 mm, pitch 1.44, voltage 120 kVp, reconstruction diameter 280 mm, exposure time 750 ms, current 60 mA, charge 45 mAs, AIDR 3D STD; and FCs were: Lung Algorithms (FC50, FC51, FC52, FC53, FC55, FC56, FC83, FC84, FC85, FC86) and Bone Algorithms (FC80, FC81, FC30, FC35).

The images were processed by the software mentioned above. The following modules were evaluated in the phantom: CTP404, for CT number linearity, CTP 528 for MTF (Modulation Transform Fourier), and CTP 486 for uniformity and noise measurement, as shown in Figure 2.

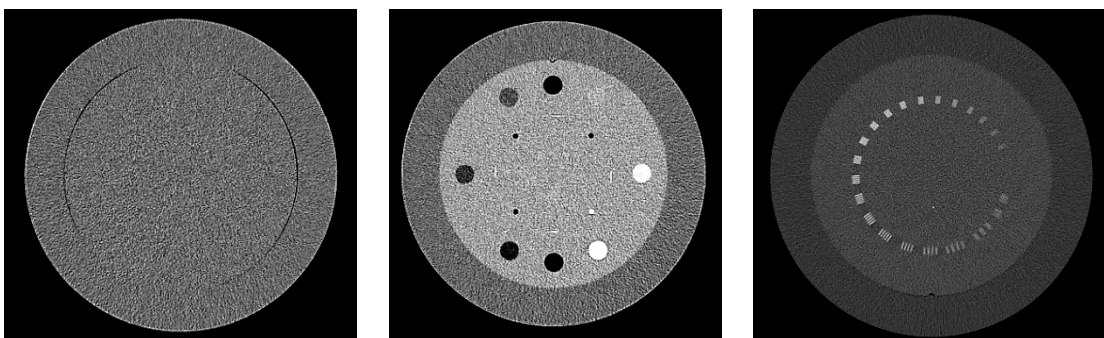
For the comparisons between the software with different FCs, the t-student statistics were performed for each test and software proposed.

2.1. Catphan® 504 phantom

There are several models and variations of the Catphan phantom for CT. The Catphan 504 phantom was developed by Varian Medical Systems, shown in Figures 1 B) and C) [4].

The Catphan 504 phantom is subdivided into modules: the CTP404 module, for CT number linearity, scan slice geometry (slice width and slice sensitivity profile), pixel size, circular symmetry, phantom position verification, patient alignment system check, scan incrementation; the CTP 528 module for measurement of high resolution with 21 line pair per cm gauge and point source, bead point source for point spread function and MTF; and the CTP 515 low contrast module with supra-slice and subslice contrast targets; and the module CTP 486 measurements of spatial uniformity, mean CT number and noise value [4].

Figure 2: Catphan 504 phantom images (from left to right): uniformity module (CTP486), sensitometry module (CTP404) and high contrast module (CTP528).



2.2. Free software for Catphan 504 phantom analysis

2.2.1. Pylinac software

Pylinac Software (V.3.1.0) was developed for the analysis of Linac tests. It has tools for automatic image analysis in the DICOM format of Catphan 504, 503 and 600, being free and open-source software. It analyses CT number linearity and image geometry (CTP404), high contrast and MTF line pairs (CTP528), HU uniformity (CTP486) and low contrast (CTP515) [10].

The software finds the Catphan modules for calculation automatically. An example of a script is shown in Figure 3. If the Regions of Interest (ROIs) in the phantom are not positioned correctly, you can modify their angles in the CTP404 module in the Pylinac source code. Additionally, the position of each module can also be adjusted as well [10].

Figure 3: Example of the script for analysing the Catphan 504 phantom [10].

```
1 from pylinac import CatPhan504
2 cbct_folder = r'catphan504_image'
3 mycbct = CatPhan504(cbct_folder)
4 mycbct.analyse()
5 print(mycbct.results())
```

2.2.2. Spice-CT software

The Spice-CT Package (V.1.8.7) is a plugin developed to be used in conjunction with the ImageJ software. This plugin enables the analysis of several quality parameters in Catphan CT images, including CT number linearity and slice thickness, uniformity and noise, helical sensitivity in the z-direction, noise between slices, noise power spectrum (NPS), modulation transfer function (MTF), low contrast, and geometry and alignment [9].

The software does not automatically choose the image slices to be calculated, it is up to the executor to select the best image for calculation.

2.2.3. CTQA-cp software

CTQA-cp (V.0.3.4) is free and open-source software. It has the analysis option for the DICOM image of Catphan 504 and 600. The analysis is automatic and it obtains the CT number linearity, slice thickness, pixel size, high resolution (MTF), image noise, noise power spectra (NPS), homogeneity, and low contrast model observer (Experimental) [11].

2.2.4. ImageJ

ImageJ is a free software, and it has tools for medical image analysis, but this software does not have automatic functions for analysing the Catphan phantom. In this sense, it will be described how to obtain the analysis of the Catphan 504 phantom [12].

2.3. Analysed image parameters

2.3.1. Uniformity

Uniformity testing can identify artefacts such as beam hardening artifacts. It may vary depending on the x-ray spectrum for imaging and phantom centring in the gantry [13-14].

Uniformity is evaluated in the uniformity module (CTP486). ImageJ was placed a ROI in the centre and four ROIs in the periphery, located at 3, 6, 9 and 12 o'clock, with a diameter of 10% of the phantom's diameter [13-14]. Uniformity was calculated for all software, with the average values of the number of CTs of their respective ROIs, by the largest difference between the average value of the central ROI and a peripheral ROI [13].

According to the IEC 61223-3-5, uniformity values for water should not be greater than ± 4 HU [14-15]. For materials other than water, IPEM suggests a level of ± 10 HU from the baseline number of other materials [15].

2.3.2. Noise

In the uniformity module (CTP486) the noise could be measured by the random variation of CT number values, and it can be estimated through the equation below. In ImageJ software, for the measurements of noise, an ROI with a diameter of approximately 40% of the diameter of the image of the phantom was obtained [1]. And the noise was obtained by equation 1:

$$N (\%) = \frac{\sigma_p}{CT_p - CT_{air}} \times 100\% \quad (1)$$

Where σ_p is the standard deviation of ROI, CT_p is the average number of ROI CTs, and CT_{air} is the mean air CT number, by definition is -1000 on the Hounsfield scale [1].

2.3.3. CT number linearity

Linearity is measured in the sensitometry module (CTP404); this test aims to evaluate the linearity of CT numbers for different materials.

In the ImageJ software, ROIs were aligned within the seven materials structures to be investigated on the image. The mean CT number for each material was obtained from this information [1, 4].

2.3.4. Slice thickness

The level value (L) was obtained to calculate the slice thickness for the analysis with the ImageJ software. To do this, the window (W) was set to its minimum value (W:0). Then, the value of L was adjusted until the ramp disappeared. The registered value of L is the maximum CT number of the ramp (CT_p). An ROI adjacent to the ramp was selected, and the background's mean CT number (CT_{bg}) was obtained. The CT number without background (CT_L) was calculated by subtracting CT_{bg} from CT_p . Next, the CT_L was divided by 2 to obtain $CT_{50\%}$. The CT number of half the maximum of the window is $CT_h =$

$CT_{50\%} + CT_{bg}$. Finally, the value of L was replaced with the value of CT_h and W was left at its minimum value (W:0) [4].

In ImageJ, a line was drawn over the ramp and the profile curve was plotted, obtaining the FWHM (full width at half maximum). Finally, the slice thickness (Z in mm) was calculated using equation 2 [4]:

$$Z(mm) = 0.42 \cdot FWHM \quad (2)$$

2.4. Python Script

A Python script was built to calculate the MTF to evaluate the high contrast resolution for comparison with the software codes.

2.4.1. High contrast resolution

To quantify the high contrast resolution in the CT image, the use of Fourier Transform Modulation (MTF) can be adopted. In module 528 there are two impulse sources (beads) aligned with the y-axis. The beads are positioned 20 mm from the centre, and 2.5 and 10 mm from the centre of the module on the z-axis [4, 16].

The bead at the top was chosen to calculate the resolution by the MTF. The MTF measurement was performed using Python code (V.3.6) and the pydicom and numpy libraries. Initially, the central image of the beads was chosen, and a square ROI around the bead was performed. Then the background noise was removed to homogenize the ROI [17].

Performed the two-dimensional Fourier transform, the values of the Fourier transform norm were summed in the y-axis direction and then normalized. The plotted curve resulted in the determination of 10% and 50% MTF.

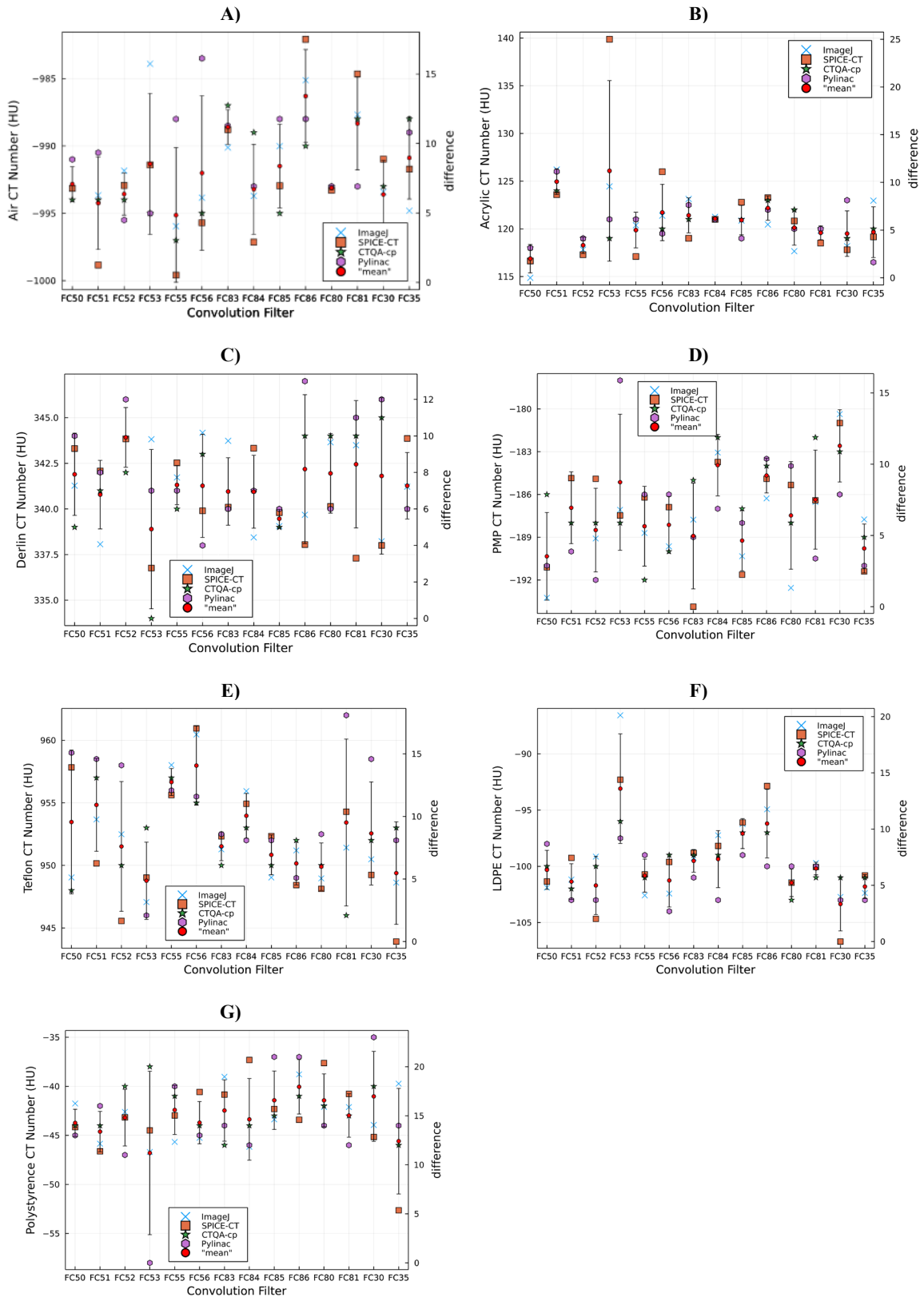
3. RESULTS AND DISCUSSIONS

The results present the data obtained by the software for each proposed test with different FCs, and thus are compared by the t-student statistics.

Figure 4 shows the values of the CT numbers for the materials air, Acrylic, Delrin, PMP, Teflon, LDPE and Polystyrene, depending on each filter chosen, for each program used. Figure 5 shows the Uniformity and Noise values. The MTF values are presented in Figure 6. Figure 7 shows the slice thickness values.

As can be seen in Figure 4, there is no coincidence in the values of the CT number between the software, but their values were in agreement with each other with $p > 0.05$. The smallest p-value among all FCs for each software and its respective test are presented in Table 1. A y-axis was added to the right of the figures to visualise the differences between the values. On this axis, the value of zero represents the minimum value of the data, and the maximum value of the axis is the greatest difference that exists between the minimum and maximum values of each figure. Uniformity data are shown in Figure 5 A). It is noted that Pylinac obtained uniformity values outside the results obtained by the other software, with a value of $p < 0.05$. And the ImageJ, Spice-CT and CTQA-cp software obtained $p > 0.05$.

Figure 4: Number of CT for A) air; B) Acrylic; C) Delrin; D) PMP; E) Teflon; F) LDPE; G) Polystyrene.



The Pylinac software does not generate the Noise result. Therefore, the Noise values were analysed only for the ImageJ, Spice-CT and CTQA-cp software, and the results coincided with each other with $p > 0.05$, as shown by the Figure 5 B).

The 50% MTF values presented in Figure 6 A). The results obtained values of $p > 0.05$ for the Python script, Pylinac, Spice-CT and CTQA-cp. And for the 10% MTF, Pylinac obtained $p < 0.05$ to filter FC52, FC83, FC84, FC85, and FC86. This significant difference may be due to the 10% MTF value returned to be an extrapolation or calculation the MTF by the Line pair per centimetre high resolution gauge method, and the Python script, Spice-CT and CTQA-cp analyse by the bead [4].

For the slice thickness values, shown in Figure 7, $p > 0.05$ was obtained for ImageJ, Spice-CT and CTQA-cp programs, and a significant difference of $p < 0.05$ for Pylinac.

Using images with different FCs made it possible to observe different results and test the limitation for data capture by the software. Figures 1 to 7 show some results of FCs outside the standard deviation range.

Figure 5: Graph of the values of A) Uniformity and B) Noise for FCs.

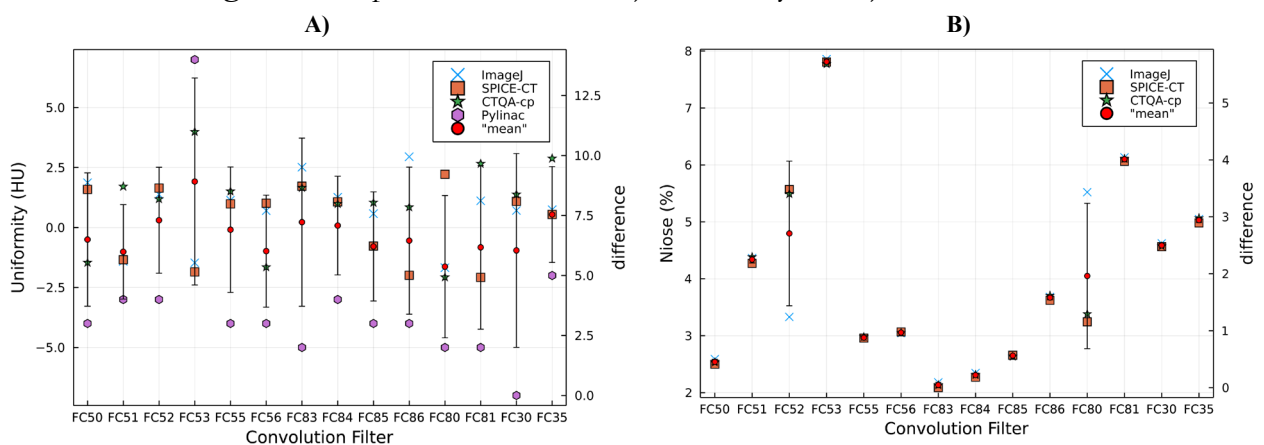


Figure 6: Graph of A) MTF 50% and B) MTF 10% by FCs.

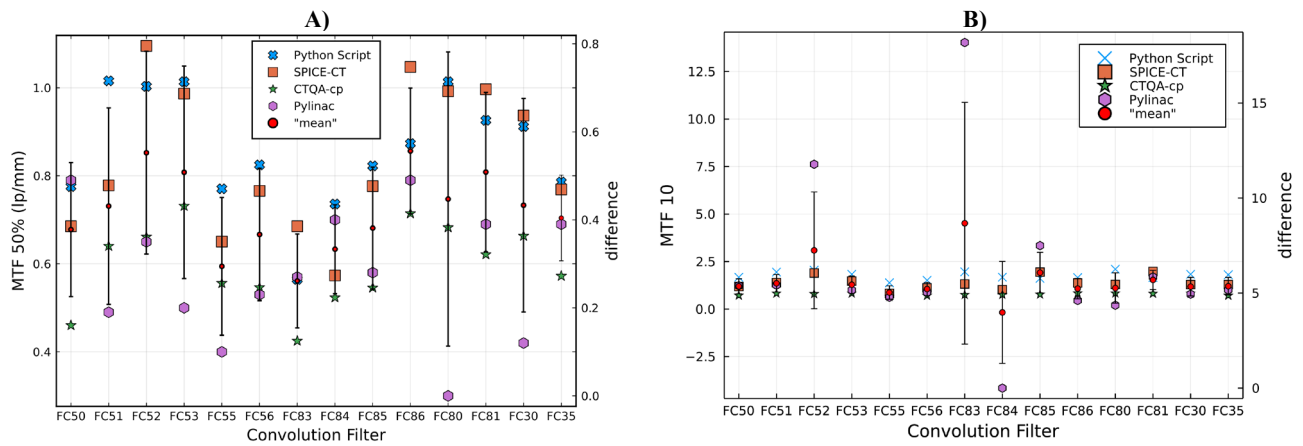


Figure 7: Slice thickness by FCs.

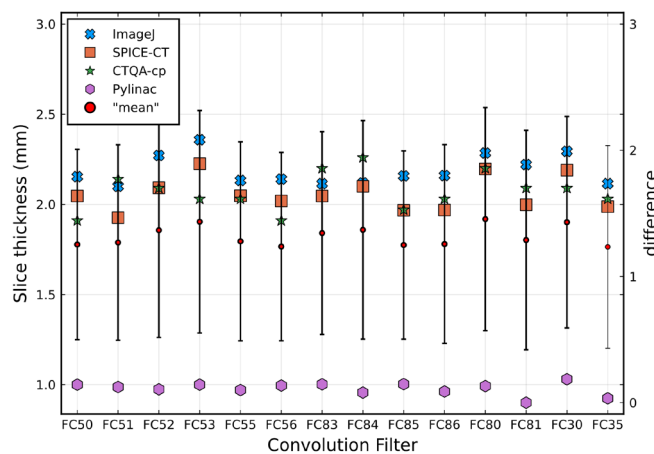





































Table 1: The smallest p-value among all FCs for each software and test evaluated with 95% confidence.

	ImageJ	Python script	Spice-CT	CTQA-cp	Pylinac
Uniformity	0.18	-	0.18	0.18	0.00
Noise	0.18	-	0.18	0.18	-
CT Number Linearity	0.06	-	0.06	0.13	0.06
Slice Thickness	0.18	-	0.19	0.19	0.00
MTF50%	-	0.08	0.08	0.06	0.08
MTF10%	-	0.11	0.20	0.11	0.00

In Figure 8, it is possible to observe the comparisons of some characteristics of the software used. The program that had the best performance in processing the tests was Spice-ct, and the one with the lowest performance was Pylinac.

Figure 8: Summary of the main features of the programs.

Software	ImageJ	Spice-CT	CTQA_cp	Pylinac
Total analysis time				
Practicality				
Viability				
Reliability				
Overall presentation of data				
Errors when loading images				
Errors in the processing of test images				
Documentation				

Legend:			
	High	Low	Medium

4. CONCLUSIONS

The use of specific software for analysing the images of the Catphan 504 phantom is important to standardize measurements and methodologies for future comparisons, and also allow a quick analysis of the CT image.

The Pylinac software does not have Catphan 504 phantom image noise analysis. It obtained a significant difference with the ImageJ, Spice-CT, and CTQA-cp programs for the uniformity and slice thickness tests; and also a significant difference for the MTF10% test with the Spice-CT, CTQA-cp, and Script Python software, which may be due to the different methodology that the software uses to calculate the MTF.

The programs ImageJ, Spice-CT, CTQA-cp, and Python Script obtained results that were consistent with each other for the uniformity, noise, CT number linearity, and slice thickness tests. For the 50% and 10% MTF tests, the results were consistent between Spice-CT, CTQA-cp, and Python Script.

It has been observed that the programs ImageJ, Spice-CT, and CTQA-cp obtained concordant results for all the tests carried out in this work. Therefore, the Pylinac software has the limitation of not calculating the standard deviation for the noise test and having a significant difference with other software for some tests, it is not recommended for use in clinical routine for the analysis of Catphan 504.

ACKNOWLEDGMENT

The authors thank Instituto Nacional de Câncer II (INCA), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, and Ministério da saúde.

FUNDING

We declare that we have received support from the Ministério da Saúde and Instituto Nacional de Câncer.

CONFLICT OF INTEREST

All authors declare that they have no conflicts of interest.

REFERENCES

- [1] International Atomic Energy Agency, Quality Assurance Programme for Computed Tomography: Diagnostic and Therapy Applications, **IAEA Human Health Series No. 19**, IAEA, Vienna, 2012.
- [2] Ministério da Saúde. Agência Nacional de Vigilância Sanitária. **Instrução Normativa - IN N° 93**, de 27 de maio de 2021.
- [3] Toshiba Medical Systems Corporation. Operation Manual Aquilion Lightning.
- [4] Phantom Laboratory Incorporated. **Catphan® 504 Manual**. 2015. Salem, NY, USA.
- [5] Tegtmeier RC, Ferris WS, Bayouth JE, Culberson WS. Performance evaluation of image reconstruction algorithms for a megavoltage computed tomography system on a helical tomotherapy unit. **Biomedical Physics & Engineering Express**, v. 8, n. 4, p.047001, 2022
- [6] Gawel J, and Kidoń J. Assessment of the CT image quality parameters on the edges of the field of view. **Physica Medica: European Journal of Medical Physics**, v.52, p.113-114, 2018.
- [7] Bellesi L, Wytenbach R, Gaudino D, Colleoni P, Pupillo F, Carrara M, et al. A simple method for low-contrast detectability, image quality and dose optimisation with CT iterative reconstruction algorithms and model observers. **European Radiology Experimental**, v.1, p.1-10, 2017

- [8] Gulliksrud K, Stokke C, and Martinsen ACT. How to measure CT image quality: variations in CT-numbers, uniformity and low contrast resolution for a CT quality assurance phantom. **Physica Medica**, v. 30, n. 4, p.521-526, 2014
- [9] Loveland, J. (2011), SPICE-CT. Available at: <https://imagej.nih.gov/ij/plugins/spice-ct/index.html>. Accessed on: 11 Nov. 2022.
- [10] Pylinac software. Available at: <https://pylinac.readthedocs.io/en/latest/cbct.html>. Accessed on: 11 Nov. 2022.
- [11] CTQA-cp software. Available at: <https://github.com/medicalphysics/ctqa-cp/releases>. Accessed on: 11 Nov. 2022.
- [12] Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA. Available at: <https://imagej.nih.gov/ij/>, 1997-2018. Accessed on: 11 Nov. 2022.
- [13] Nowik P, Bujila R, Poludniowski G, Fransson A. Quality control of CT systems by automated monitoring of key performance indicators: a two-year study. **J Appl Clin Med Phys**, v. 16, p.254–265, 2015.
- [14] Institute of Physics and Engineering in Medicine. Recommended Standards for the Routine Performance Testing of Diagnostic X-ray Imaging Systems, **IPEM Report 91**. York: IPEM, p.1–112, 2010.
- [15] Gulliksrud, K., Stokke, C., & Martinsen, A. C. T. How to measure CT image quality: variations in CT-numbers, uniformity and low contrast resolution for a CT quality assurance phantom. **Physica Medica**, v.30, n.4, p.521-526, 2014.
- [16] Kayugawa A, Ohkubo M, Wada S. Accurate determination of CT point-spread-function with high precision. **J Appl Clin Med Phys**, v.14, n. 4, p.3905, 2013.
- [17] Hak, E. Z., Anam, C., Budi, W. S., & Dougherty, G. An improvement in automatic MTF measurement in CT images using an edge of the PMMA phantom. En **Journal of Physics: Conference Series**. IOP Publishing, v. 1505, n. 1, p. 012039, 2020.
- [18] Mail TB. Catphan® 700 Product Guide. Phantom Laboratory Incorporated. 2021.

LICENSE

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third-party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material.

To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.