Ther Adv Cardiovasc Dis
2020, Vol. 14: 1–6
DOI: 10.1177/1753944720924254

Inter-observer variation of Syntax score among cardiac surgeons, clinical and interventional cardiologists

Gustavo Neves de Araujo, Tiago Luiz Luz Leiria, Mariana Vargas Furtado, Bruno da Silva Matte, Guilherme Pinheiro Machado, Ana Maria Krepsky, Luiz Carlos Corsetti Bergoli, Sandro Cadaval Goncalves, Marco Vugman Wainstein and Carisi Anne Polanczyk

Abstract

Background: Despite the complexity of SYNTAX score (SS), guidelines recommend this tool to help choosing between coronary artery bypass grafting (CABG) and percutaneous coronary intervention (PCI) in patients with left main of three-vessel coronary artery disease. The aim of this study was to compare the inter-observer variation in SS performed by clinical cardiologists (CC), interventional cardiologists (IC), and cardiac surgeons (CS).

Methods: Seven coronary angiographies from patients with left main and/or three-vessel disease chosen by a heart team were analyzed by 10 CC, 10 IC and 10 CS. SS was calculated *via* SYNTAX website.

Results: Kappa concordance was very low between CC and CS (k = 0.176), moderate between CS and IC (k = 0.563), and moderate between CC and IC (0.553). There was a statistically significant difference between CC, who classified more cases as low complexity (70%), and CS, who classified more cases as moderate complexity (80%) ($p = 0.041$).

Conclusion: Concordance between SS analyzed by CC, CS and IC is low. The usefulness of SS in decision-making of revascularization strategy is undeniable and evidence supports its use. However, this study highlights the importance of well-trained professionals on calculating the SS. It could avoid misclassification of borderline cases.

Keywords: coronary artery bypass graft, left main, multivessel disease, percutaneous coronary intervention, SYNTAX score

Received: 27 April 2018; revised manuscript accepted: 15 April 2020.

Correspondence to:
Gustavo Neves de Araujo
Department of Cardiology, Hospital de Clinicas de Porto Alegre, Rua Ramiro Barcelos, 2350, Porto Alegre, RS 90035-903, Brazil
Faculty of Medicine, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
gustavon.araujo@gmail.com

Tiago Luiz Luz Leiria
Mariana Vargas Furtado
Bruno da Silva Matte
Ana Maria Krepsky
Department of Cardiology, Hospital de Clinicas de Porto Alegre, Porto Alegre, Brazil

Guilherme Pinheiro Machado
Luiz Carlos Corsetti Bergoli
Sandro Cadaval Goncalves
Marco Vugman Wainstein
Carisi Anne Polanczyk
Faculty of Medicine, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
Department of Cardiology, Hospital de Clinicas de Porto Alegre, Porto Alegre, Brazil

Introduction

Studies comparing coronary artery disease treatment with coronary artery bypass grafting (CABG) and percutaneous coronary intervention (PCI) showed superiority of CABG mainly in prevention of further interventions,[1,2] especially in patients with anatomically complex lesions. With the continuous development of new techniques and devices used in PCI, it became necessary to compare CABG and PCI in patients with more severe disease. The SYNTAX study compared the two treatment methods in patients with left main or three-vessel disease.[3] A score created to grade the anatomical complexity of coronary lesions showed that patients with less severe disease could be treated with both CABG and PCI, with similar results.

The 2014 European Society of Cardiology/European Association for Cardio-Thoracic Surgery (ESC/EACTS) guidelines for myocardial revascularization included for the first time the SYNTAX score (SS) to help choose the intervention method.[4] Despite its unquestionable usefulness,[5–7] there may be some degree of variability among different physicians in the interpretation of angiographies, due to the

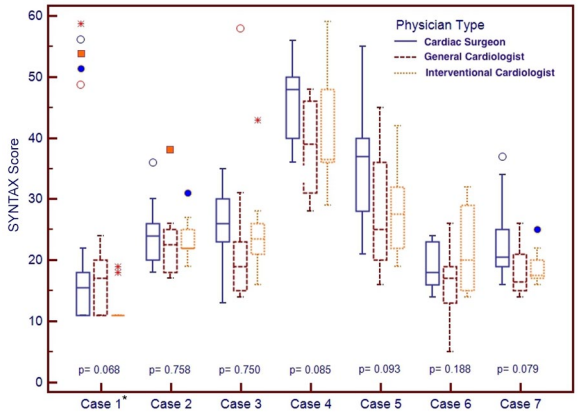

Figure 1. Mean/Median* SYNTAX score among different cardiology professionals.

presence of some subjectivity in scoring. Interpretations can differ between more or less conservative professionals, more or less experienced professionals, and even by the same professionals at different times.[8,9] The aim of this study was to compare the inter-observer variation in SS calculation performed by clinical cardiologists (CC), interventional cardiologists (IC), and cardiac surgeons (CS).

Methods

We conducted a cross-sectional study among CS, CC, and IC. All were recognized specialists in their medical societies, and all worked in tertiary care/teaching hospitals in southern Brazil. In all, 94 specialists were invited to participate either personally or by email, and 30 completed the study.

We chose seven coronary angiographies from patients with left main and/or three-vessel disease, chosen by a heart team comprising a CS, a CC and an IC. The same heart team analyzed the films together to determine the gold standard score for each film. None of this group participated as research subjects in the study. We created a website where we uploaded the coronary angiographies, so the research subjects could analyze the films from anywhere they could use a computer. After analyzing each film, they were redirected to the SYNTAX calculator website,[10] and the results were sent directly to our database.

This project was submitted and approved by the local Ethics Committee. Informed consent was filled at the moment the physician logged into the website. There was no violation of confidentiality of data regarding the patients whose angiography was analyzed.

Statistical analysis

SS was analyzed as both continuous and categorized variables, according to the original article, where low complexity was a score from 0 to 22, medium complexity was a score from 23 to 32, and high complexity was a score above 32. Variables were expressed as mean and standard deviation (SD), or median and interquartile range (IQR). One-way ANOVA was performed to determine the difference between means of syntax score among the three groups. Variation between groups and reproducibility of SS was determined by Kappa index. Fisher's exact test was used in the analysis of contingency tables. Bimodal p value <0.05 is considered to indicate statistical significance. Data were analyzed using SPSS (version 18.0.0; IBM Company, Chicago, IL, USA) and MedCalc (version 12.5.0.0, bvba , Mariakerke, Belgium).

Sample size was calculated based on an inter-observer and intra-observer concordance of 0.45 and 0.59, respectively.[8] Assuming a Kappa coefficient of 0.75 between the evaluators, we estimated that a sample of 294 evaluations would be enough to determine the correlation coefficient with a power of 80% and level of significance of 5%.

Results

We included 10 CC, 10 CS, and 10 IC. When we analyzed SYNTAX as a continuous variable, mean SS was not different among different specialists (Figure 1). Mean SS among different professionals are presented in Table 1.

Kappa concordance was very low between CC and CS ($k=0.176$), moderate between CS and IC ($k=0.563$), and moderate between CC and IC (0.553; Table 2).

There was a statistically significant difference between CC, who classified more cases as low complexity (70%), and CS, who classified more cases as moderate complexity (80%; $p=0.041$). Agreement rates of categorized syntax score calculated by different specialists compared with the gold standard were as low as 10% and as high as 100%, depending on specialist group and coronary angiography (Table 3).

Table 1. Syntax score among different professionals and Heart Team's Gold Standard.

Health care professional/ case number		1	2	3	4	5	6	7
Gold standard (heart team)		11	20	29	41	39	16	23
CC	Mean (SD)	16.7 (4.54)	22.7 (6.32)	22.9 (13.35)	38.1 (7.59)	27.4 (9.6)	16.6 (6.13)	18.3 (4.27)
	p	0.003	0.210	0.182	0.258	0.004	0.764	0.007
	Mean difference	5.7	2.7	– 6.1	–2.9	–11.6	0.6	–4.7
CS	Mean (SD)	15.4 (4.27)	24.4 (5.4)	26.1 (5.89)	45.9 (6.38)	35.9 (9.66)	18.9 (3.87)	23.1 (7.03)
	p	0.010	0.030	0.154	0.038	0.337	0.420	0.965
	Mean difference	4.4	4.4	– 2.9	4.9	–3.1	2.9	0.1
IC	Mean (SD)	12.5 (3.17)	23.3 (3.55)	24.8 (7.37)	40.5 (8.89)	28.6 (7.67)	21.5 (6.98)	18.7 (2.9)
	p	0.169	0.019	0.105	0.862	0.002	0.034	0.001
	Mean difference	1.5	3.2	– 4.2	–0.5	–10.4	5.5	–4.3

CC, clinical cardiologist; CS, cardiac surgeon; IC, interventional cardiologist; SD, standard deviation.

Discussion

The present study indicates that, in patients with left main and/or three-vessel disease, concordance between SS analyzed by CC, CS and IC is from moderate to very low. Moreover, IC tend to evaluate cases as being of lower severity. This study is, to our knowledge, the first study to compare the variability of SS calculated by different cardiology specialists.

The SYNTAX trial represented a change of paradigm in coronary artery disease, by permitting patients with complex coronary anatomy to choose between different treatment strategies depending on their complexity score. This finding was so striking that the 2014 ESC/EATS guidelines for myocardial revascularization included SS as a method to help choose the intervention method.[4]

Table 2. Mean/median* Syntax score and (risk category) by case.

	CC *n* = 10	CS *n* = 10	IC *n* = 10
Case 1*	17 (Low)	15.5 (Low)	11 (Low)
Case 2	22.7 (Low)	24.4 (Moderate)	23.2 (Moderate)
Case 3	22.9 (Low)	26.1 (Moderate)	24.8 (Moderate)
Case 4	38.1 (High)	45.9 (High)	40.5 (High)
Case 5	27.4 (Moderate)	45.9 (High)	28.2 (Moderate)
Case 6	16.6 (Low)	18.9 (Low)	21.2 (Low)
Case 7	18.3 (Low)	23.1 (Moderate)	18.8 (Low)

Kappa index: CC *versus* CS = 0.176; IC ($p = 0.392$) *versus* CS = 0.563 ($p = 0.032$); CC *versus* IC = 0.533 ($p = 0.024$).
CC, clinical cardiologist; CS, cardiac surgeon; IC, interventional cardiologist; SD, standard deviation.

Accurate coronary anatomy determination is essential, and SS was created to make the decision process as objective as possible. However, a certain degree of subjectivity exists, and characteristics such as conservativeness and experience in interpretation angiographies may change the final result of analysis. Généreux *et al.* found poor inter-observer agreement (k = 0.33) between three IC who analyzed 30 multivessel angiograms, with a substantial improvement (k = 0.76) after

Table 3. Agreement rate in categorized Syntax score among specialists.

Test	Categorized SYNTAX score (Gold Standard)	Agreement rate		
		Clinical	Surgeon	Interventional
1	Low	90%	100%	100%
2	Low	50%	40%	70%
3	Moderate	20%	80%	40%
4	High	60%	100%	90%
5	High	30%	60%	20%
6	Low	80%	70%	50%
7	Moderate	20%	10%	10%

advanced training.[11] Two other studies analyzing inter-observer variability between two IC showed only moderate agreement (k = 0.56 and 0.58).[12,13]

The aim of our study was to compare analysis of SS mainly by CS and IC. Our hypothesis was that CS could be less conservative and overestimate the result; on the other hand, IC could be more conservative due to possible professional bias. The CC would be a group to balance the other two described above. Interventional bias is likely the manifestation of two well-recognized forms of bias: self-interest bias and confirmation bias.[14] This is a reality of modern medicine, and recognition is the first step towards overcoming it.

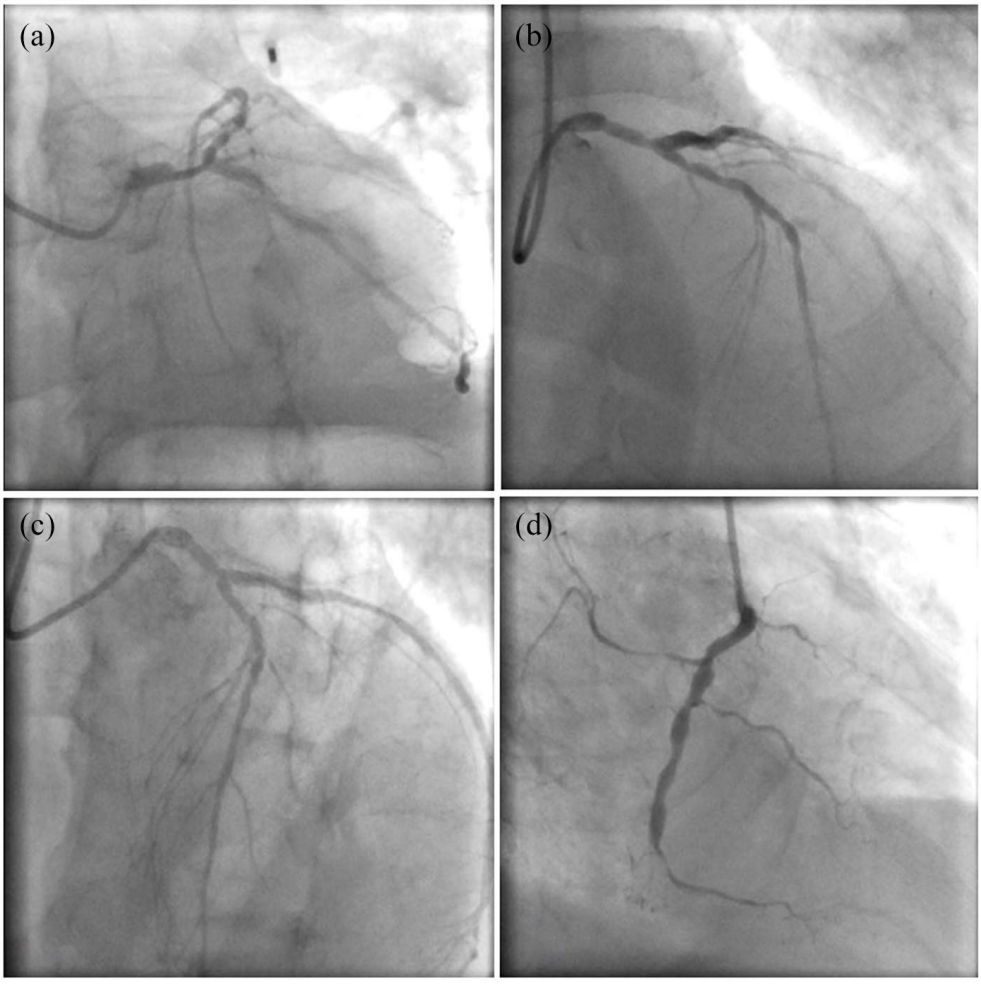

Figure 2. Case 5 coronary angiography images. (a) Right anterior oblique caudal projection of left coronary artery. (b) Posteroanterior cranial projection of left coronary artery. (c) Left anterior oblique cranial projection of left coronary artery. (d) Left anterior oblique projection of right coronary artery.

We did not find the differences we were expecting. Mean SS was not statistically different between groups, although there was a trend to significance in some cases analyzed. This may be explained by the reduced sample of 10 observers in each group, characterizing a Beta error. Kappa concordance was from moderate to poor. Moreover, there was a clear difference between and within groups regarding SS categorical classification as low, moderate, and high disease severity. In case 5 (Figure 2), for example, the average score of CC and IC classified SS as moderate, which would make angioplasty acceptable. CS, however, classified SS as high, which would prohibit PCI in this case. This indicates that SS may be a valuable tool for the global orientation of the severity of the lesions in a particular patient, but the final decision should depend on a Heart Team decision, as well as local expertise (treating more complex left main stenosis, for example) and patient conditions.

A guideline recommendation of such an important decision needs external validity, and our findings corroborate the literature showing poor agreement of the score in a real-world situation. Two recent studies with a very similar methodology compared CABG and PCI in patients with left main coronary disease, and have shown opposite results.[15,16] Several theories have been proposed to explain such differences, and different SS calculation may be one of them.

Limitations

The main limitation of our study is the reduced sample of observers in each group, although we have reached the calculated sample size. SS is not a user-friendly tool to determine coronary complexity, making enrollment a difficult job. However, we do bring the message that SS has a poor reproducibility, and recommendations regarding its use in treatment decisions should be made with caution.

Conclusion

In conclusion, SS concordance between CC, CS, and IC is low. The usefulness of SS in decision-making of coronary revascularization strategy is undeniable and evidence supports its use. However, this study highlights the importance of well-trained professionals in calculating the SS, which could avoid misclassification of borderline cases.

Approval number

2010-0044

Conflict of interest statement

The authors declare that there is no conflict of interest.

Ethics Committee

Comitê de Ética em Pesquisa do Hospital de Clínicas de Porto Alegre (CEP-HCPA).

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by grants from the Fundo de Incentivo à Pesquisa do Hospital de Clínicas de Porto Alegre (FIPE-HCPA), Porto Alegre, Brazil.

ORCID iD

Gustavo Neves de Araujo https://orcid.org/0000-0002-1989-1670

References

1. Serruys PW, Unger F, Sousa JE, *et al.* Comparison of coronary-artery bypass surgery and stenting for the treatment of multivessel disease. *N Engl J Med* 2001; 344: 1117–1124.
2. The SoS Investigators. Coronary artery bypass surgery versus percutaneous coronary intervention with stent implantation in patients with multivessel coronary artery disease (the Stent or Surgery trial): a randomised controlled trial. *Lancet* 2018; 360: 965–970.
3. Serruys PW, Morice M-C, Kappetein AP, *et al.* Percutaneous coronary intervention versus coronary-artery bypass grafting for severe coronary artery disease. *N Engl J Med* 2009; 360: 961–972.
4. Kolh P, Windecker S, Alfonso F, *et al.* 2014 ESC/EACTS Guidelines on myocardial revascularization: the Task Force on Myocardial Revascularization of the European Society of Cardiology (ESC) and the European Association for Cardio-Thoracic Surgery (EACTS). Developed with the special contribution of the European Association of Percutaneous Cardiovascular Interventions (EAPCI). *Eur J Cardiothorac Surg* 2014; 46: 517–592.

5. Capodanno D, Capranzano P, Di Salvo ME, *et al*. Usefulness of SYNTAX score to select patients with left main coronary artery disease to be treated with coronary artery bypass graft. *JACC Cardiovasc Interv* 2009; 2: 731–738.

6. Valgimigli M, Serruys PW, Tsuchida K, *et al*. Cyphering the complexity of coronary artery disease using the syntax score to predict clinical outcome in patients with three-vessel lumen obstruction undergoing percutaneous coronary intervention. *Am J Cardiol* 2007; 99: 1072–1081.

7. van Gaal WJ, Ponnuthurai FA, Selvanayagam J, *et al*. The Syntax score predicts peri-procedural myocardial necrosis during percutaneous coronary intervention. *Int J Cardiol* 2009; 135: 60–65.

8. Serruys PW, Onuma Y, Garg S, *et al*. Assessment of the SYNTAX score in the Syntax study. *EuroIntervention* 2009; 5: 50–56.

9. Herrman JP, Azar A, Umans VA, *et al*. Inter- and intra-observer variability in the qualitative categorization of coronary angiograms. *Int J Card Imaging* 1996; 12: 21–30.

10. SYNTAX SCORE. SYNTAX Score website, http://www.syntaxscore.com/calc/syntaxscore/frameset.htm (accessed 10 February 2018).

11. Genereux P, Palmerini T, Caixeta A, *et al*. SYNTAX score reproducibility and variability between interventional cardiologists, core laboratory technicians, and quantitative coronary measurements. *Circ Cardiovasc Interv* 2011; 4: 553–561.

12. Tanboga IH, Ekinci M, Isik T, *et al*. Reproducibility of syntax score: from core lab to real world. *J Interv Cardiol* 2011; 24: 302–306.

13. Garg S, Girasis C, Sarno G, *et al*. The SYNTAX score revisited: a reassessment of the SYNTAX score reproducibility. *Catheter Cardiovasc Interv* 2010; 75: 946–952.

14. Andrew J., Foy AJ and Filippone EJ. The case for intervention bias in the practice of medicine. *Yale J Biol Med* 2013; 86: 271–280.

15. Stone GW, Sabik JF, Serruys PW, *et al*. Everolimus-eluting stents or bypass surgery for left main coronary artery disease. *N Engl J Med* 2016; 375: 2223–2235.

16. Makikallio T, Holm NR, Lindsay M, *et al*. Percutaneous coronary angioplasty versus coronary artery bypass grafting in treatment of unprotected left main stenosis (NOBLE): a prospective, randomised, open-label, non-inferiority trial. *Lancet* 2016; 388: 2743–2752.