

Molecular Docking and Analysis of *In Silico* Generated Ligands against SARS-CoV-2 Spike and Replicase Proteins

Ifeanyichukwu Okeke*, Cosmas Okeke

Department of Surgery, Ahmadu Bello University Teaching Hospital, Kaduna State, Nigeria

ABSTRACT

By way of molecular docking, a library of *in silico* generated ligands was docked to SARS-CoV-2 spike and replicase proteins to identify leads with propensity to bind them with high affinity. The identified leads proved to bind these proteins with stronger affinity than the native ligand aiding in their *in vivo* metabolic processes. Whereby it was observed that spike protein binds to its cellular receptor with binding affinity of -4.8 Kcal/mol; it binds to a non-cellular analogue with -5.4, while 4twy 3BL and 5n19 D03 bind spike protein with binding affinities of -7.3 Kcal/mol each. They also bind replicase protein with -8.2 and -7.2 Kcal/mol respectively. The findings indicate that the identified ligands can preferentially displace or inhibit binding of the viral proteins to their native endogenous ligands and that both cellular attachment through spike and ACE2 interaction, and viral replication process can both be inhibited by using just one of the substances identified. From the study, 5c8s G3A and 2d2d ENB were identified as the most suitable leads that are favorably disposed for SARS-CoV-2 spike protein detection from biological samples, while 3d62 959 and 1r4l XX5 were identified as leads with most suitable drug likeness against SARS-CoV-2 based on the filters from SwissADME and Molinspiration cheminformatics and therefore deserve further *in vitro* and *in vivo* evaluations.

Keywords: *In silico*; Molecular docking; COVID19; SARS-CoV-2; Drug likeness

INTRODUCTION

The outbreak of the novel coronavirus also known as coronavirus disease 2019 (COVID19), or Severe Acute Respiratory Syndrome Corona Virus 2 (SARS-CoV-2) in the latter part of the year 2019 took the entire human race unawares with its devastating health, social and economic consequences. Since the outbreak, scientists all over the world have swung into action, researching in various disciplines to find solutions to contain the virus. Among the factors limiting the effort in containing the virus especially in underdeveloped and developing countries of the world are the issues of diagnosis and treatment [1]. Effective diagnosis and isolation of infected persons is one of the key non pharmaceutical means to contain the spread of the virus. Currently, the only valid test for COVID19 is the Nucleic Acid Amplification Test (NAAT), such as Real-Time Reverse-Transcription Polymerase Chain Reaction (rRT-PCR). This test is not readily available and affordable

especially to the people in the under developed and developing countries of the world, hence making access to COVID19 test difficult in these areas. To effectively control and manage any emerging, reemerging and novel diseases, early detection and characterization is very important.

Molecular interactions between proteins driving activities of SARS-CoV-2 and exogenous smaller molecules (ligands) which are able to bring about structural modification following occupation of the binding sites of the viral proteins thereby altering their biological functions are considered in this study as a way to enhance research in the aspects of diagnosis and treatment for the novel disease. Findings have shown that the interaction between ligands and the protein receptors which induce conformational changes in the proteins can bring about alteration in the thermodynamic stability of proteins rather than their mechanical stability with formation of new complexes that affect the activities of that particular protein [2]. These smaller

Correspondence to: Ifeanyichukwu Okeke, Department of Surgery, Ahmadu Bello University Teaching Hospital, Kaduna State, Nigeria, E-mail: arnyic@yahoo.com

Received: 03-Jan-2023, Manuscript No. JACR-23-20576; **Editor assigned:** 06-Jan-2023, PreQC No. JACR-23-20576 (PQ); **Reviewed:** 20-Jan-2023, QC No. JACR-23-20576; **Revised:** 27-Jan-2023, Manuscript No. JACR-23-20576 (R); **Published:** 03-Feb-2023, DOI: 10.35248/2155-6148.23.14.1093

Citation: Okeke I, Okeke C (2023) Molecular Docking and Analysis of *In Silico* Generated Ligands against SARS-CoV-2 Spike and Replicase Proteins. J Anesth Clin Res. 14: 1093

Copyright: © 2023 Okeke I, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

molecules or ligands also known as drugs tend to find their ways into the binding pockets of the receptors molecule. The interaction which can be likened to an enzyme-inhibitor interaction which can be a competitive one whereby a native ligand is inhibited from binding to the receptor in preference for the exogenous one depending on their binding affinities and conditions that enhance it. This particular concept of molecular interaction can be utilized to identify lead compounds which can be applied in laboratory detection and treatment of COVID-19.

From the foregoing, two gene products from SARS-CoV-2: the spike protein and the replicase protein were docked to their ligands as generated from the PDB (rcsb.org) to identify leads that can serve as diagnostic and therapeutic agents directly or after some possible structural optimization. The outcome of the findings indicate that most of the ligands can bind the replicase protein thereby inhibiting viral replication, while many others can bind spike protein thereby inhibiting viral attachment to their receptors in the host cell. Some of the ligands also have the ability to bind both spike and replicase proteins with much greater binding affinity than that with which the cell receptors bind them [3,4].

In doing justicetothis study,weapplied the reliability to the *in Silico* prediction that has been widely applied in preliminary drug discovery strategy to, (1) Dock a library of computer generated ligands to SARS-CoV-2 spike and replicase proteins in other to identify leads with maximum hits with the proteins. Analyze the interacting residues of the proteins with the best leads identified. (2) Study drug suitability of the identified ligands with the best hit. The interactions observed could also be exploited in the study of suitable biochemical diagnostic alternative to SARS-CoV-2 or COVID19 Nucleic Acid Amplification Test (NAAT), such as real-time reverse-transcription polymerase chain reaction (rRT-PCR).

SARS-CoV-2 is a highly contagious respiratory pathogen infecting humans with a very rapid spreading rate, causing most common symptoms such as fever, dry cough and tiredness, and less common symptoms such as aches and pains, sore throat, diarrhea, conjunctivitis, headache, loss of taste or smell and skin rash or discoloration of fingers or toes [5]. Serious symptoms may include difficulty breathing or shortness of breath, chest pain or pressure, loss of speech or movement. A recent report from John Hopkins University indicates that latest global total cases as at January 15, 2021 stood at 93,620,509 confirmed cases with daily new cases of 746,642.

The mucus membrane lining the nose, mouth and eye until lately are the entry points for the virus into the vascular networks and subsequent invasion of cells of various organs such as the nervous system lungs, heart and many others. Once into the vascular networks, entry of COVID19 into the erythrocytes is mediated by anion exchange membrane band3 protein in its tetrameric molecular structure. SARS-CoV-2 structurally has four proteins: Spike (S), Membrane (M), Envelope (E) and Nucleocapsid (N) proteins. The spike protein is responsible for attachment to the host's cellular receptor. Here we considered one structural (the spike protein) and one nonstructural

(replicase) proteins in studying the interaction of the viral proteins with some of their ligands.

Attachment and fusion of SARS-CoV-2 to the host's cell receptor-Angiotensin Converting Enzyme 2 (ACE 2) through a series of process that will not be illustrated here is facilitated by the spike protein. The expression of the S protein in some coronaviruses by an infected cell can also mediate fusion of the infected cell with adjacent uninfected cells leading to formation of multinucleated cells or syncytia. This has been considered as a strategy to allow viruses to spread directly between cells thereby avoiding virus-neutralizing antibodies.

The Envelope protein (E) is an integral membrane protein involved in many activities that have to do with the life cycle of the virus such as budding, packaging, envelope formation and pathogenesis. It also functions as an ion-channeling viroporin and facilitates interactions between other coronavirus proteins and host cell proteins. The E protein is richly expressed inside the infected cell during replication cycle but only a little amount of it is incorporated into virion envelope while majority are deposited at the point of intracellular trafficking such as the endoplasmic reticulum, Golgi apparatus and ERGIC where it carries out the functions of viral packaging and budding.

The membrane protein (M) is the most abundant of all the structural proteins. It is regarded as the central organizer because of its role in interacting with the rest of the coronavirus structural proteins to bring about formation of virion envelope and stabilization of the N protein-RNA complex (nucleocapsid) which ultimately promotes completion of viral assembly. The primary mechanisms that direct coronavirus RNA synthesis and processing are situated within the nonstructural proteins nsp7 to nsp16. These are cleavage products of two large replicase polyproteins translated from the coronavirus genome. The Nucleocapsid (N) has the ability to bind coronavirus RNA genome to constitute the nucleocapsid Apart from being involved with processes relating to viral genome, the N protein is also involved in coronavirus replication activity and response of the host cell to infection by the virus. The replicase, though an accessory and Non-Structural Protein (nsp), is indispensable in the replication cycle of the N protein (a structural protein). The resolution of the structure of the replicase protein (nsp9) suggests that the protein comprises a single β -barrel with a fold previously unseen in single domain proteins [6]. The fold superficially resembles an OB-fold with a C-terminal extension and is related to both of the two subdomains of SARS-CoV 3C-like protease (which belongs to the serine protease superfamily). nsp9 has, presumably, evolved from a protease. The crystal structure suggests that the protein is dimeric.

MATERIALS AND METHODS

The Protein Data Bank (PDB) structures of the receptors used: spike (PDB ID: 2ghv) and replicase (PDB ID: 1uw7 (nsp9)) proteins in their prepared forms using Discovery Studio visualize v20.1.0.19295 is shown in Figures 1a and 1b. Structures of some of the ligand ions used are shown in figure. This work was carried out on windows 8.1 Pro with processor: Intel® Core (TM) i5 CPU M 520 @ 2.40GHz having installed memory

(RAM): 4.00GB (3.86 GB usable) on system type: 64 bit operating system although 32-bit Windows Vista operating system also work well. PyRx docking software version 0.8 for Windows (<http://pyrx.sourceforge.net>) was used for molecular docking. PyRx is open source software to perform virtual screening [7]. It is a combination of several softwares such as AutoDock Vina, AutoDock 4.2, Mayavi, Open Babel, etc. PyRx uses Vina and AutoDock 4.2 as docking softwares. In this study, AutoDock Vina was used. Discovery studio (Discovery Studio: v20.1.0.19295), Pymol (Pymol stereo 3D quad buffer) and ICM Browser (Molsoft MolBrowser 3.8-7d) were used to examine structural properties and study binding interactions between receptor residues and the ligands. The chemical structures of the receptors (2ghv and 1uw7 (nsp9)) and those of their ligands were downloaded from protein data bank (rcsb.org) and PubChem (<https://pubchem.ncbi.nlm.nih.gov/>). Canonical SMILES and other information about the ligands and the receptor were extracted from PubChem. The structures of the receptors spike and replicase proteins were retrieved by searching in the Protein Data Bank (PDB).

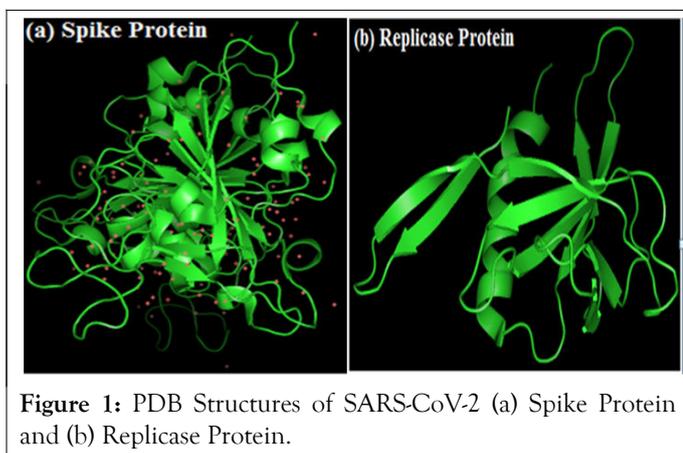


Figure 1: PDB Structures of SARS-CoV-2 (a) Spike Protein and (b) Replicase Protein.

Macromolecule methods

The x-ray structures of the receptors (2ghv and 1uw7 (nsp9)) were downloaded from protein data bank, hetatoms were removed using discovery studio and resaved (Figure 1a). The input ligand files were also prepared for virtual screening by minimizing their energies and converting them to PDBQT file format when they were imported into PyRx software as chemical table file. Auto dock Vina took each ligand and bonded its different conformations to the macromolecules (2ghv and 1uw7 (nsp9)) separately to get the binding energies in different orientations of each ligand [8]. Docking was repeated three times on the same system specifications for the purpose of process validation and all returned minimal variation ($P < 0.01$ data not shown) in uff energy, binding energy and RMSD values in the two receptors. Discovery studio (Discovery Studio: v20.1.0.19295), Pymol (Pymol stereo 3D quad buffer) and ICM Browser (Molsoft MolBrowser 3.8-7d) were all used to visualize and analyze binding interactions between residues of the receptor molecules and the ligands [9,10].

RESULTS

A library of the ligands belonging to the SARS-CoV-2 spike and replicase proteins were docked to their binding pockets in the receptors. Compounds in the library demonstrated good binding affinity with many having higher binding affinity to the spike protein than Angiotensin Converting Enzyme 2 (ACE2), its native (cellular) ligand. Analysis was restricted to 10 structures with binding energy -4.8 Kcal/mol and lower. The highest binding affinity recorded in the screening is -7.3 and -8.2 Kcal/mol for spike and replicase proteins respectively with 1r4l NAG (ACE2) as the reference molecule for comparing extent of interaction between spike protein and ligands. Given the same ligand, replicase protein generally shows higher binding affinity than spike protein implying higher tendency to interact with the ligands [11].

Here are the various types of interactions between the ligands and the residues of the spike protein receptor within the binding pocket [12]. It was observed that the interactions are predominantly non covalent in nature typical of those for maintaining 3D structures of large molecules such as proteins and nucleic acids. These interactions suggest that both the ligands and the receptor are large biological molecules which specifically but transiently bind to one another to bring about biological and pharmacological responses as the case may be and these are relatively weak electric forces that attract neutral molecules. These forces can be seen in the complex formed between molecules of 4twy 3BL, 5n19 D03, 5c8s G3A, 1r4l XX5, 3d62 959 and 1r4l NAG and the receptor residues. Different schools of thought have different explanation for this type of interaction [13]. While some believe that this type of interaction arose because the model is not a perfect imitation of the real structure, others believe that they are due to unstable and maximum torsional strain. Melesina Jelena opined that unfavorable interactions in molecular docking do not necessarily mean that the compounds are not good inhibitors. This is considered the most favorable (or rewarded contact) with the highest level of complementarities between binding site and ligand [14,15]. An interaction whereby different kinds of heavy atom contacts are rewarded as seen in interaction between 5c8s G3A, 2alv CY6, 2d2d ENB and 1r4l NAG with the receptor. Carbon-Hydrogen Bond is common among organic compounds. It is a form of covalent interaction where a carbon shares its outer valence electron with up to four hydrogen atoms. This bond confers more strength to the complex formed than van der Waals forces [16].

DISCUSSION

The presence of conventional hydrogen bond in cellular ACE 2 (1r4l NAG) can confer adequate resistance against displacement by any ligand of high binding affinity without a hydrogen bond suitably in the same position of amino acid residue as seen in ACE2. The presence of a conventional hydrogen bond therefore can pose a resistance in attempt to displace viral spike protein from cell surface except in situations where exogenous ligands adequately matched it with enough suitable bonds as we have in 5c8s G3A and 2d2d ENB [17]. The amino acids that dominated interaction between spike and ACE2 in the binding pocket

constitute the most important targets in any interaction aimed at interfering with binding of ACE2 to spike protein. Considering the ten selected ligands, ligands with amino acids matching most of those involved in interaction between ACE2, and possessing high binding affinity will be favorably disposed to elicit maximum activity against the viral spike protein [18].

1r4l NAG (ACE2) exhibited much lower binding affinity to spike

Ligand	Unfavorable (UI)	Interaction	Favorable Interaction (FI)	Total	% FI
1r4l NAG	3		6	9	66.7
3r24 SAM	6		9	15	60
2gx4 NOL	8		11	19	57.9
4twy 3BL	7		7	14	50
2d2d ENB	10		9	19	47.4
2alv CY6	9		8	17	47.1
3d62 959	5		4	9	44.4
5c8s G3A	11		8	19	42.1
5n19 D03	10		7	17	41.2
1wof I12	13		9	22	40.9
1r4l XX5	9		6	15	40

Table 1: Quantization of favorable and unfavorable interactions among ligands.

Efficient complex formation between ligands and receptors can also be utilized as diagnostic makers when they yield distinctive characteristic reactions such as precipitation, colour change, gas production or change in temperature following the interaction between the ligand and receptors. A detection method could also be electrochemical or fluorescence, aimed at detecting the presence of target molecules in biological samples that are known to bind to the receptor and the amount of precipitate, colour, gas or temperature change determines the level of the viral protein present, implying that the assay can be qualitative and quantitative [19].

In vivo activities of replicase protein 1uw7 (nsp9) from this study is likely to be disrupted by quite a good number of ligands as many ligands demonstrated high binding affinity with the receptor. Of the many ligands that exhibited high binding affinity with the two receptors, only those whose binding affinities in the first orientation ranged from -7.3 to -4.8 and -8.2 to -5.0 Kcal/mol for 2 ghv and 1uw7 respectively were selected for analysis [20]. It could be observed that auto dock structures appear to have different modification from their sdf counterparts seen in Figure 2. Target analysis of some of the selected ligands was done only for ligands which showed high level of binding

protein than many of the spike protein ligands. This implies that spike protein can be inhibited from binding to ACE2 or ACE2 can easily be displaced from binding pocket in the presence of any of the molecules which have higher binding affinity thereby disrupting cell invasion by the virus. The highest binding affinity attained by ACE2 is -4.8 Kcal/mol whereas a spike protein ligand 4twy 3BL has -7.3 Kcal/mol (Tables 1 and 2).

affinity, including ACE2 to the receptors. Visualization of binding pockets, nature and strength of interaction between ligands and receptor residues were the parameters considered.

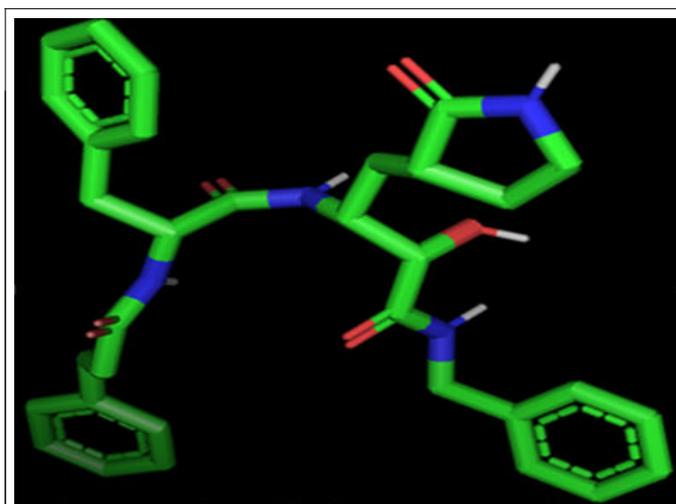


Figure 2: Autodock Structure of 5n19 Do3 in Association with replicase.

S/N	Molecule	SwissADMET cLogP	Molinspiration LogP	Mean LogP
1	4TWY 3BL	3.97	4.6	4.29
2	5N19 D03	2.66	2.39	2.53
3	5C8S G3A	-5.47	-4.67	-5.1
4	2ALV CY6	3.07	2.67	2.87
5	2D2D ENB	2.57	4.46	3.52
6	2GX4 NOL	3.14	5.04	4.1
7	1WOF I12	1.94	1.1	1.52
8	3R24 SAM	-2.96	-4.14	-3.55
9	3D62 959	1.46	1.09	1.28
10	1R4L XX5	2.06	0.28	1.17

Table 2: Comparison of cLogP and miLogP values for theselected compounds as seen in SwissADMET and Molinspiration platforms respectively.

Among the selected ligands, only 3d62 959 and 1r4l XX5 exhibited complete drug like properties based on Lipinski, Veber and Ghose rules. 3d62 959 has 40% and 1r4l XX5 has 50% of its residue matching those of ACE2 in the receptor binding pocket, with no hydrogen bonds. 3d62 959 also have many of its residues forming unfavorable bonds and CYS C: 323, ASN C: 330 and PRO C: 324 not forming any bond at all. All these indicate lower competitiveness among the rest of the ligands against ACE2 for spike protein receptor.

CONCLUSION

In this study, an extensive study was carried out *in silico* to discover molecules that could be of medical use for management of COVID-19 in the areas of diagnosis and treatment. Ligands (or small molecules or drugs) with activity against SARS-CoV-2 spike and replicase proteins were extracted from designated drugs and proteins databases. The ligands were screened against the receptors (SARS-CoV-2 spike and replicase proteins) to isolate molecules that can favorably bind these receptors thus displacing or inhibiting binding of their endogenous or native ligands-Acetyl Choline Esterase 2 (ACE2) in the case of spike protein to inhibit viral attachment to the host cells. Following the analysis and observations so far, all the 10 selected compounds having higher binding affinity with viral spike protein than ACE2 deserve further *in vitro* and *in vivo* evaluations to determine which compound has the best *in vivo* activity against SARS-CoV-2 spike and replicase proteins, and safe for human use and also efficient in detecting viral proteins from biological samples with high precision.

RECOMMENDATIONS

This study therefore recommends repurposing the use of 5c8s

G3A, 2d2d ENB, 3d62 959 and 1r4l XX5 in further studies of COVID19 treatment and possible alternative diagnostic approach.

AUTHOR'S CONTRIBUTIONS

Ifeanyichukwu Okeke conceptualized and coordinated the study. Ifeanyichukwu Okeke and Cosmas Okeke drafted the manuscript, Cosmas Okeke assisted with the molecular visualization, toxicity analyses and figures preparation.

FINANCIAL SUPPORT

NIL

CONFLICTS OF INTEREST

NIL

REFERENCES

1. Walls AC, Park YJ, Tortorici AM, Wall A, McGuire AT, Velesler D. Structure Function and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*. 2020; 181(2): 281-292.
2. Lonie JA, Tierney KM, Ebmeier KP .Screening for mild cognitive impairment: a systematic review. *Int. J. Geriatr. Psychiatr.* 2019; 24(2): 902-915.
3. Deckers K, Van BMP, Schiepers OJ, De VM, Muñoz Sánchez JL, Anstey KJ, Kate I, et al. Target risk factors for dementia prevention: a systematic review and Delphi consensus study on the evidence from observational studies. *Int J Geriatr Psychiatr.* 2015; 30(3): 234-246.
4. Hafkemeijer A, Altmann-Schneider I, Slagboom PE, Van der Grond J, Rombouts SA. Associations between age and gray matter volume in anatomical brain networks in middle-aged to older adults. *Aging Cell.* 2005; 13(6): 1068-1074.
5. Zakzanis KK, Mraz R, Graham SJ. An fMRI study of the Trail Making Test. *Neuropsychologia.* 2005; 43(13): 1878-1886.

6. Enzinger C, Fazekas F, Matthews PM, Ropele S, Schmidt H, Smith S, et al. The Risk factors for progression of brain atrophy in aging: six-year follow-up of normal subjects. *Neurology*. 2005; 64(1): 1704-1711.
7. Messerotti BS, Patron E, Zanatta P, Polesel E, Palomba D, Scott A. Pre-existing cognitive status is associated with reduced behavioral functional capacity in patients 3 months after cardiac surgery: an extension study. *Psychiatr*. 2014; 36(4): 368-374.
8. Silbert BS, Scott DA, Evered LA, Lewis MS, Maruff PT. Pre-existing cognitive impairment in patients scheduled for elective coronary artery bypass graft surgery. *Anesthesia & Analgesia*. 2007; 104(5): 1023-1028.
9. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005; 53(4): 695-699.
10. Arbuthnott K, Frank J. Trail making test, part B as a measure of executive control: validation using a set-switching paradigm. *J Clin Exp Neuropsychol*. 2000; 22(4): 518-528.
11. Llinàs-Reglà J, Vilalta-Franch J, López-Pousa S, Calvó-Perxas L, Torrents Rodas D, Garre-Olmo J. The trail making test. *Assessment*. 2017; 24(2): 183-196.
12. Giovagnoli AR, Del Pesce M, Mascheroni S, Simoncelli M, Laiacona M, Capitani E. Trail making test: normative values from 287 normal adult controls. *Ital J Neurol Sci*. 1996; 17(4): 305-309.
13. Messerotti Benvenuti S, Patron E, Zanatta P, Polesel E, Palomba D. Pre-existing cognitive status is associated with reduced behavioral functional capacity in patients 3 months after cardiac surgery: an extension study. *Psychiatr*. 2014; 36(4): 368-374.
14. Stern Y. Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurol*. 2012; 11(11): 1006-1012.
15. Ninomiya T. Diabetes mellitus and dementia. *Curr Diabet Rep*. 2014; 14(5): 482-487.
16. Pires P W, Dams Ramos CM, Matin N, Dorrance AM. The effects of hypertension on the cerebral circulation. *Am J Physiol Heart Circ Physiol*. 2013; 304(12): 1598-1614.
17. Meguro K, Dodge HH. Vascular mild cognitive impairment: identifying disease in community-dwelling older adults, reducing risk factors, and providing support. *J. Alzheimer's Dis*. 2019; 70(S1): 293-302.
18. Gustafson D. Adiposity indices and dementia. *Lancet Neurol*. 2006; 5(8): 713-720.
19. Thompson PM, Moussai J, Zohoori S, Goldkorn A, Khan AA, Mega MS, et al. Cortical variability and asymmetry in normal aging and Alzheimer's disease. *Cerebr Cortex*. 1998; 8(6): 492-509.
20. Demarin V, Zavoreo I, Kes VB. Carotid artery disease and cognitive impairment. *J Neurol Sci*. 2012; 322(2): 107-111.