

**Supplementary Table 1** Methodological quality assessment using the QUADOMICS Tool.

<b>Ref.</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>score</b>
Cabrerizo (2017) <sup>20</sup>	Y	N/A	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	Y	12/14
Buisset (2019) <sup>21</sup>	Y	N/A	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	Y	12/14
Myer (2020) <sup>22</sup>	Y	N/A	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	Y	11/14
Pan (2020) <sup>23</sup>	Y	N/A	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	Y	11/14
Burgess (2015) <sup>24</sup>	Y	N/A	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	N	11/14
Leruez (2018) <sup>25</sup>	Y	N/A	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	Y	11/14
Kouassi Nzougnet (2019) <sup>26</sup>	N	N/A	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	Y	11/14
Javadiyan (2012) <sup>27</sup>	Y	N/A	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	Y	12/14
Umeno (2019) <sup>17</sup>	Y	N/A	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	Y	12/14
Gong (2020) <sup>28</sup>	Y	N/A	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	N/A	11/13
Rossi (2019) <sup>29</sup>	N	N/A	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	N/A	10/13
Rong (2017) <sup>32</sup>	Y	N/A	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	N	11/14
Chen (2019) <sup>33</sup>	Y	N/A	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	N	10/14
Myer (2020) <sup>34</sup>	Y	N/A	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	Y	11/14
Barbosa Breda (2020) <sup>35</sup>	Y	N/A	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N/A	N	Y	12/14

Item 1. Were selection criteria clearly described? 2. Was the spectrum of patients representative of patients who will receive the test in practice? 3. Was the type of sample fully described? 4. Were the procedures and timing of biological sample collection with respect to clinical factors described with enough detail? 4a. Clinical and physiological factors 4b. Diagnostic and treatment procedures 5. Were handling and pre-analytical procedures reported in sufficient detail and similar for the whole sample? And, if differences in procedures were reported, was their effect on the results assessed? 6. Is the time period between the reference standard and the index test short enough to reasonably guarantee that the target condition did not change between the two tests? 7. Is the reference standard likely to correctly classify the target condition? 8. Did the whole sample or a random selection of the sample receive verification using a reference standard of diagnosis? 9. Did patients receive the same reference standard regardless of the result of the index test? 10. Was the execution of the index test described in sufficient detail to permit replication of the test? 11. Was the execution of the reference standard described in sufficient detail to permit its replication? 12. Were the index test results interpreted without knowledge of the results of the reference standard? 13. Were the reference standard results interpreted without knowledge of the results of the index test? 14. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice? 15. Were uninterpretable / intermediate test results reported? 16. Is it likely that the presence of over-fitting was avoided; Y=criteria achieved, N=criteria not achieved, N/A=not applicable.