

**Supplementary 3. Newcastle-Ottawa Scale**

Study	A. Selection			B. Comparability	C. Outcome		Score			
	Representativeness of exposed cohort	Selection of non-exposed cohort	Ascertainment of exposure		Demonstration that outcome of interest was not present at start of the study	Comparability of cohorts on the basis of the design or analysis		Assessment of outcome	Was follow-up long enough for outcomes to occur	Adequacy of follow up of cohorts
<b>A. Magnetic - anchor</b>										
Gotoda et al. (2009) <sup>19</sup>	*	*	*	*	-	*	*	*	*	7
Matsuzaki et al. (2018) <sup>44</sup>	*	-	*	*	-	*	*	-	-	4
Ye et al. (2019) <sup>49</sup>	*	*	*	*	**	*	*	*	*	9
<b>B. Double endoscope</b>										
Uraoka et al. (2010) <sup>22</sup>	*	*	*	*	**	*	*	*	*	9
Higuchi et al. (2013) <sup>28</sup>	*	*	*	*	**	*	*	*	*	9
Ogata et al. (2017) <sup>40</sup>	*	-	*	*	-	*	*	*	*	6
Çolak et al. (2019) <sup>48</sup>	*	-	*	*	-	*	*	-	-	4
<b>C. Clip involving methods</b>										
Okamoto et al. (2012) <sup>24</sup>	*	*	*	*	**	*	*	-	-	7
Okamoto et al. (2012) <sup>25</sup>	*	*	*	*	**	*	*	-	-	7
Ota et al. (2012) <sup>26</sup>	*	*	*	*	**	*	*	-	-	7
Matsumoto et al. (2013) <sup>29</sup>	*	*	*	*	**	*	*	-	-	7
Cai et al. (2015) <sup>31</sup>	*	-	*	*	-	*	*	*	*	6
Suzuki et al. (2016) <sup>34</sup>	*	*	*	*	**	*	*	-	-	7
Yoshida et al. (2016) <sup>37</sup>	*	*	*	*	**	*	*	-	-	7
Yamada et al. (2016) <sup>35</sup>	*	*	*	*	**	*	*	-	-	7
Yamasaki et al. (2016) <sup>36</sup>	*	-	*	*	-	*	*	-	-	4
Noda et al. (2016) <sup>33</sup>	*	*	*	*	**	*	*	-	-	7
Xie et al. (2017) <sup>41</sup>	*	*	*	*	**	*	*	-	-	7
Jacques et al. (2017) <sup>38</sup>	*	-	*	*	-	*	*	-	-	4
Kitagawa et al. (2018) <sup>43</sup>	*	-	*	*	-	*	*	*	*	6
Hashimoto et al. (2018) <sup>42</sup>	*	*	*	*	**	*	*	-	-	7
Zhang et al. (2019) <sup>47</sup>	*	-	*	*	-	*	*	*	*	6
Zhang et al. (2019) <sup>50</sup>	*	-	*	*	-	*	*	*	*	6

Study	A. Selection			Demonstration that outcome of interest was not present at start of the study	B. Comparability		C. Outcome		Score
	Representativeness of exposed cohort	Selection of non-exposed cohort	Ascertainment of exposure		Comparability of cohorts on the basis of the design or analysis	Assessment of outcome	Was follow-up long enough for outcomes to occur	Adequacy of follow up of cohorts	
D. Miscellaneous methods									
Imaeda et al. (2009) <sup>20</sup>	*	-	*	*	-	*	-	-	4
Yonezawa et al. (2006) <sup>18</sup>	*	*	*	*	**	*	*	*	9
Hijikata et al. (2012) <sup>23</sup>	*	*	*	*	-	*	-	-	5
Motohashi et al. (2009) <sup>21</sup>	*	-	*	*	-	*	-	-	4

According to the scale's "star system", each study receives a score based upon three different perspectives: study selection groups; comparability; and ascertainment of outcome of interest.  
A. Selection

1. Representativeness of intervention cohort—a] Selection from general population or general hospital\*; b] only selected group of patients; c] no description of inclusion/exclusion criteria
2. Selection of non-intervention cohort—a] drawn from same community as intervention cohort\*; b] drawn from different source; c] no description of the derivation of the non-intervention cohort
3. Ascertainment of outcome—a] examination protocol or department's archive\*; b] no description
4. Demonstration that outcome was not present at start of study—a] yes\*; b] no

#### B. Comparability

1. Comparability of cohorts on basis of design or analysis—a] study controls for participants' age and gender\*; b] study controls for examinations' indication\*

#### C. Outcome

1. Assessment of outcome—a] independent physician's assessment\*; b] record linkage\*; c] self-report; d] no description
2. Was follow up long—a] yes\*; b] no

Adequacy of follow up of cohort—a] complete follow up\*; b] minimal loss to follow up ( $\leq 20\%$ ) all ages included, all diseases, or description of those lost suggesting no difference from those followed\*; c] follow up rate  $< 80\%$  and no description of losses to follow up or description suggesting differences from those followed; d] no statement.