Introduction

Patients with a suspect of pancreatic solid lesion require a cyto-histological diagnosis which is obtained in most cases by means of endoscopic ultrasound (EUS) [1]. The optimal means to obtain EUS-guided biopitic material is debated both in terms of which needle to use and of the optimal sampling technique [2]. The two most widely employed sampling techniques are “suction” (SU) and “slow-pull” (SP). Suction takes advantage of the negative pressure exerted by a syringe, usually 10 or 20 mL, in which vacuum is created, connected to the needle [3]. Slow-pull consists in the slow removal of the stylet creating a minimal negative pressure inside the needle, which is estimated to be around 5% of the force generated with the syringe [4]. SP has the advantage of less contamination with blood and therefore of a possible increased diagnostic yield. The two techniques are similarly popular, as suggested by the results of a recent survey on the attitude of 186 endoscopic ultrasonographers from the United States, Europe, and Asia, which showed that 47% prefer to employ SU and 42% SP [5].
European Society of Gastrointestinal Endoscopy (ESGE) guidelines suggest using SU “applying a continuous suction with a syringe” for sampling solid pancreatic masses [6]. This statement is, however, mostly limited to fine-needle aspiration (FNA) needles and based on a limited number of randomized controlled trials (RCTs) that demonstrate that 10 mL of suction during sampling with 22G or 25G FNA needles improves accuracy compared with standard non-suction technique [7].

Moreover, although some retrospective studies had suggested that SP is not inferior to SU [8], at the time of publication of the ESGE guidelines there was no sufficient evidence to recommend for or against using the stylet SP technique.

However, more recently, after publication of ESGE guidelines, several non-randomized and randomized trials have been performed comparing SU and SP [9–11] in sampling solid pancreatic lesions, trying to clarify this aspect.

We, therefore, aimed to perform a meta-analysis of RCTs to compare adequacy, sensitivity, specificity, accuracy of the SU and SP techniques for sampling of solid pancreatic lesions.

**Methods**

**Search strategy and inclusion/exclusion criteria**

The Pubmed database was searched from inception to May 30, 2019 for the following search terms: (suction OR slow-pull OR slowpull OR slow pull OR tissue acquisition OR FNA OR FN B OR fine needle) AND (EUS OR endosonography OR endoscopic ultrasound OR endoscopic ultrasonography) AND (pancreas OR pancreatic). Also, the abstracts of the following meetings were searched for other potentially includable studies: Digestive Disease Week (DDW) 2016–2018, United European Gastroenterology Week (UEGW) 2016–2018, European Society of Digestive Endoscopy Days (ESGE Days) 2017–2018.

The inclusion criteria were: 1) EUS being performed for the presence of a pancreatic mass; 2) the study design being a RCT comparing the SU and SP techniques (no matter which needle was employed); and 3) at least one of the following as outcome measures: diagnostic adequacy, diagnostic accuracy, sensitivity, specificity, adverse events. The exclusion criteria were: 1) study design being a review, case-report, animal model study or any non-retrospective or prospective comparison not being a RCT and 2) language other than English. All the retrieved titles were screened to exclude articles not related to the study aim (screening process). The remaining studies were further screened by abstract and full-text analysis to identify those that fulfilled all inclusion and exclusion criteria (eligibility process). References in the included studies were also screened to find potential additional studies. Reasons for exclusion were recorded for each study. In the case of cohort overlap, the most recent and/or the largest cohort was included; in case of uncertainty, the authors of the studies were contacted for clarification. The screening process was independently performed by L.A. and G.C., with disagreement resolved by consensus or by discussion with the third author (M.C.P.). The methodology was developed from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist [12].

**Definitions**

We accepted the definitions of SU and SP provided by in the individual studies if it was clarified that “suction” was applied using a syringe before moving the needle within the lesion with to-and-fro movements and “slow-pull” when it was reported that the stylet was slowly removed once lesion penetrated over with to-and-fro movements of the needle within the lesion, without applying any suction.

**Data extraction**

Data extraction was independently performed by L.A. and G.C. and disagreement resolved by consensus or by discussion with the third author (M.C.P.). The following variables were extracted in a dedicated datasheet for each included study: first author, Journal Title, publication year, study design and primary outcome, details on randomization, power calculation (yes/no), crossover (yes/no), blinding between endoscopist and pathologist (yes/no), Study setting (unicenter or multicenter), country of origin, accrual period, number of cases, gender distribution, mean (median) age of patients, sampling procedure type (endoscopic ultrasound-guided fine-needle aspiration [EUS-FNA] or endoscopic ultrasound-guided fine-needle biopsy [EUS-FNB]), needle size, endoscopic instrument, mean (median) mass size, provided definition of suction and slow-pull method, other procedure variables (fanning, ROSE), post-procedure monitoring for side-effects, gold-standard for diagnosis, technical success for SU and SP, diagnostic accuracy for SU and SP, diagnostic accuracy for SU and SP, mean number of needle passes required for diagnosis for SU and SP, quantification of blood contamination for SU and SP, sensitivity and specificity for SU and SP, adverse events of SU and SP. The quality of included studies was assessed by one author (M.C.P.) and discussed with a second one (G.C.) using the Cochrane risk of bias tool for Randomized Controlled Trials [13] for assessing the risk of bias, with the following 7 domains: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcomes assessment, incomplete outcome data, and selective reports, other bias.

As suggested, the study quality was considered poor if one criterion was not met (i.e. high risk of bias for one domain) or two criteria were deemed unclear and the assessment that this was likely to have biased results, fair if one criterion was not met (i.e. high risk of bias for one domain) or two criteria were deemed unclear and the assessment that this was unlikely to have biased results. Finally, quality was considered good if all criteria were met. As we aimed at distinguishing studies with a high risk of bias, we analyzed separately studies with good or fair quality as compared with those with poor quality.

**Statistical analysis**

A meta-analysis of all eligible studies identified was carried-out with the software package Comprehensive Meta-Analysis (Bio-Stat, Englewood, New Jersey, United States) using a random-effects model [14, 15]. In addition to within-study variance, the random-effects model takes in consideration heterogeneity among studies.
and gives more conservative estimates. We, therefore, present results of the random-effects model as we believe that the relevant variation in the risk is most likely a consequence of inter-study differences. The quantity of heterogeneity was assessed by means of the I² value [16]. We considered an I² value of 25 % or lower as trivial heterogeneity, and an I² value of 75 % or higher as considerable heterogeneity. Publication bias was assessed using the Begg and Mazumdar test. Events were expressed as proportions and 95 % confidence intervals (CI) were calculated using exact methods and assuming a Poisson distribution. Odds ratios (OR) with 95 % confidence intervals (CI) were used to report the risk estimates. P<0.05 was considered statistically significant. Before the analysis, we developed the following a priori hypotheses to examine their effect on the performance of SU and SP and to explore reasons for any heterogeneity: study quality (either poor or fair/good), Geographic Region of origin (either Asia or Western Countries), type of comparison (same lesion with different techniques in a randomized order or distinct lesions randomized to distinct techniques), presence of ROSE (yes/no), type of needle (FNB vs FNA needles).

Results

Search results and study characteristics

Overall, 2,906 potential studies were identified throughout the initial Pubmed search, and three additional studies were identified by screening of conferences’ proceedings; 2,862 of these 2,909 studies were excluded after screening of the title and abstract, as they were not related to the study topic or did not fulfill the inclusion criteria. Forty-seven studies seemed to meet the inclusion criteria and their full-text were screened and considered for the analysis, of which 40 were excluded, because these did not report the necessary data of interest, they met some exclusion criteria, or they were duplicate publications. Seven studies eventually remained for qualitative analysis and quantitative synthesis, (Fig.1).

Table 1 summarizes the characteristics of the seven included studies [10, 11, 17–21]. They were all full-text RCTs published between 2017 and 2019; two were initially available only as abstracts from conferences’ proceedings at the time of submission that during the revision process became available as full text and whose data were accordingly updated [20, 21]. Only two studies [10, 20] were multicentric. Two studies were carried out in the US [10, 11], 2 in Korea [18, 19] and one each in Brazil [21], India [17] and Italy [20]. Different needles were employed, with 3 studies employing FNB needles [11, 19, 20] and 4 FNA [10, 17, 18, 21]). Four of the seven studies randomized distinct individual patients with solid pancreatic lesions to either SU or SP sampling [10, 11, 17, 20], for a total of 327 patients (164 SP, 163 SU); the remaining 3 studies [18, 19, 21] randomized the same patients with solid pancreatic lesions to both SU and SP sampling in a random order, for a total of 148 patients. Cytopathology evaluation with ROSE was available in only one study [10].

The study quality was fair/good in all studies but one (assessed as poor [18].

Diagnostic adequacy

Four of the seven studies [11, 17–19] reported data on the diagnostic adequacy comparing SU and SP. Samples were considered adequate in case of “cytological aspirates or histological core tissue samples that were sufficient for diagnostic interpretation” [11], or if “adequate for the cytopathologist to make a diagnosis” [17], or if “sufficient amount of representative cells for pathological diagnosis” [18] was provided, while the definition was not clearly provided in one study [19]. There was no significant difference between the two techniques (Fig.2) with an OR of 0.98 (95 % CI 0.44–2.17; P=0.96) and no heterogeneity (I²=0 %). There seemed to be no indication of publication bias for this analysis (Begg and Mazumdar test, Kendall’s tau P=0.08).

When considering only the three studies with fair/good quality scores, the results were similar to those of the pooled analysis with an OR=0.86 (95 % CI 0.38–1.97) and no heterogeneity (I²=0 %), while in the study considered to have a lower quality score the adequacy seemed higher with SU (OR=5.21, 95 %
### Table 1  Characteristics of the included studies.

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Country</th>
<th>Study setting and accrual period</th>
<th>Needle (manufacturer)</th>
<th>Pancreatic lesions</th>
<th>Gender (M/F)</th>
<th>Age (mean±SD)</th>
<th>Lesion size (mm), mean ± SD (range)</th>
<th>Design</th>
<th>ROSE</th>
<th>Study quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saxena P</td>
<td>2017</td>
<td>USA</td>
<td>Multicenter Sep 2013–Dec 2014</td>
<td>22G (Expect Slimline; Boston Scientific), FNA</td>
<td>121 61 SP 60 SU</td>
<td>70 M/51 F</td>
<td>64 ± 13.8</td>
<td>32 ± 1.6 28 ± 1.2</td>
<td>Randomization of different patients to one arm</td>
<td>Yes</td>
<td>Fair/good</td>
</tr>
<tr>
<td>Bansal RK</td>
<td>2017</td>
<td>India</td>
<td>Single center Jan 2015–Dec 2015</td>
<td>22-gauge needle (Echo-Tip Ultra HD; Cook) FNA</td>
<td>36 of which 18 SP 18 SU</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Randomization of different patients to one arm</td>
<td>No</td>
<td>Fair/good</td>
</tr>
<tr>
<td>Weston BR</td>
<td>2017</td>
<td>USA</td>
<td>Single center Feb 2012–Jul 2013</td>
<td>22G and 25G (ProCore; Cook Medical) FNB</td>
<td>60 30 SP 30 SU</td>
<td>35 M/25 F</td>
<td>31 (14 70) 34 (10–67)</td>
<td>Randomization of different patients to one arm</td>
<td>No</td>
<td>Fair/good</td>
<td></td>
</tr>
<tr>
<td>Lee JM</td>
<td>2018</td>
<td>Korea</td>
<td>Single center Aug 2015–Jul 2016</td>
<td>22G (Expect Slimline; Boston Scientific) FNA</td>
<td>48</td>
<td>29 M/19 F</td>
<td>68.1 ± 11.9</td>
<td>37.6 ± 16.7</td>
<td>Randomization of same patients to both arms in different order</td>
<td>No</td>
<td>Poor</td>
</tr>
<tr>
<td>Cheng S</td>
<td>2019</td>
<td>Brazil</td>
<td>Single center May 2015–June 2016</td>
<td>22G (Expect Slimline; Boston Scientific or Medi-Globe GmbH)</td>
<td>50</td>
<td>29 M/21 F</td>
<td>63.9 ± 10.4</td>
<td>42/50 (84 %) &gt;40 mm</td>
<td>Randomization of same patients to both arms in different order</td>
<td>No</td>
<td>Fair/good</td>
</tr>
<tr>
<td>Di Mitri R</td>
<td>2019</td>
<td>Italy</td>
<td>Multicenter NR</td>
<td>20G (ProCore, Cook Medical) FNB</td>
<td>110 55 SP 55 SU</td>
<td>49 M/61 F</td>
<td>70.9 ± 11.3</td>
<td>35.9 mm ± 19.9 in SP and 37.6 ± 14.1 in SU</td>
<td>Randomization of different patients to one arm</td>
<td>No</td>
<td>Fair/good</td>
</tr>
<tr>
<td>Lee KY</td>
<td>2019</td>
<td>Korea</td>
<td>Single center Aug 2015–Jul 2016</td>
<td>22G (ProCore, Cook Medical) FNB</td>
<td>50</td>
<td>27 M/23 F</td>
<td>66.8 ± 12.6</td>
<td>33.2</td>
<td>Randomization of same patients to both arms in different order</td>
<td>No</td>
<td>Fair/good</td>
</tr>
</tbody>
</table>

FNA, fine-needle aspiration; FNB, fine-needle biopsy; SP, slow-pull; SU, suction
95% CI 0.24–111.54, I² = 0%), although not significantly. As for the Geographic Region of origin, 3 of the studies included in this analysis were conducted in Asia [17, 19] and do not suggest an advantage of one of the two techniques (OR 1.00 95% CI 0.301–3.366, I² = 15.8), as also for the single study performed in the US OR (1.2 95% CI 0.322–4.469, I² = 0%). In the two studies that randomized distinct individual patients with solid pancreatic lesions to either SU or SP sampling [11,17], there was no difference between the 2 techniques (OR = 1.35 95% CI 0.42–4.34) with no heterogeneity (I² = 0%); similar results were also observed in the 2 studies sampling the same solid pancreatic lesions with both SU and SP sampling in a random order (OR = 1.06; 95% CI 0.14–7.84) with a slightly higher heterogeneity (I² = 44.2%). None of these studies employed ROSE. Finally, the employed needles did not affect the adequacy of SU and SP, as both in the 2 studies using a 22G ProCore needle [11, 19] (OR=0.77; 95% CI 0.32–1.86; I² = 0%), and in the 2 studies using 25G FNA needles [17, 18] (OR = 3.04 95% CI 0.43–21; I² = 0%) there was no difference in the adequacy between SP and SU with no heterogeneity.

### High degree of blood contamination

Five of the seven examined studies [17, 21] reported data on the degree of blood contamination. We calculated the rate of highest degree of blood contamination as defined in the 5 studies as “blood clots present” [17], “blood in >50% of the slide” [19], “interference with the making of pathological diagnosis due to large amount of blood cells” [18], significant blood contamination [20] and red blood cells >50% [21].

The pooled rate of high blood contamination resulted higher with SU, being of 19.7% (95% CI 6.8%–45.4%; I² = 92.6%) with SP and 27.6% (95% CI 13.2%–48.8%; I² = 90%) with SU (Supplementary Fig. 1a and Supplementary Fig. 1b). All these studies did not employ ROSE. In the studies with fair-good quality scores the pooled rate of high blood contamination was 21% with SP and 30.6% with SU. The difference remained in the two studies conducted on distinct individual patients with solid pancreatic lesions [17,20], with a pooled rate of high blood contamination of 13.6% for SP and 22% for SU, while it resulted less evident being 24.7% for SP and of 26.4% (95% CI 5.9%–24.7%, I² = 42.8) for SU in the 3 studies sampling the same lesions in a random order [18,19,21]. Notably, while in the 2 studies employing FNB [19,20] the pooled rate of high blood contamination was 13.7% with SP and 15.7% with SU, in the 3 studies using FNA [17,18,21] this was 25.2% with SP and 37.4% with SU.

### Diagnostic accuracy, sensitivity and specificity

Four of the seven studies [10,18,20,21] reported sufficient data to compare the diagnostic accuracy of SU and SP. We first calculated the pooled rate of diagnostic accuracy with the 2 techniques that resulted to be very similar being 89.5% (95% CI 80.1%–94.7%; I² = 60.5%) with SP and 84.6% (95% CI 67.8%–93.5%; I² = 82%) with SU (Supplementary Fig. 2a and Supplementary Fig. 2b), suggesting increased heterogeneity with SU.

However, there was no significant difference between the two techniques (Fig.3) with an OR of 0.82 (95% CI 0.36–1.85; P = 0.63) and moderate heterogeneity (I² = 52.4%). There was no publication bias affecting this analysis (Begg and Mazumdar test, Kendall’s tau P = 0.73). These results were not affected by the study design, with the OR being 0.86 (95% CI 0.26–2.8) and moderate heterogeneity (I² = 41.7%) in the studies sampling distinct patients in the 2 arms [10, 20], and 0.79 (95% CI 0.15–4.04), though with considerable heterogeneity (I² = 78.1%), in the 2 studies sampling the same lesions [18, 21]. Also, the difference between SP and SU was not significant in the single study using ROSE [10] and in the 3 not using it [18, 20,21] (OR 1.01; 0.29–3.42; P = 0.98; I² = 64.4%). As for the Geographic region of origin of the studies, the only study originating from Asia [18] showed an accuracy significantly favoring SP (OR = 0.34; 95% CI 0.12–0.99; P = 0.05) that was not appreciated in the three remaining studies (OR = 1.10; 95% CI 0.45–2.69; P = 0.82; I² = 43.7%). As for study quality, the only study graded as with poor quality suggested a superiority of SP [18] (see Fig.3). Finally, three studies employed FNA needles [10, 18, 21] and reported no superiority of SP or SU (OR = 0.69; 95% CI 0.28–1.70; P = 0.42; I² = 58.7%), as also the single study with FNB [20].
The pooled sensitivities were 88.7% (95% CI 80.3%–93.8%) with moderate heterogeneity (I² = 51.2%) with SP and 83.4% (95% CI 64.6%–93.2%) with considerable heterogeneity (I² = 84.8%) with SU, respectively (Supplementary Fig. 3a and Supplementary Fig. 3b). Pooled rates of specificity were 97.2% (95% CI 78.6%–99.7%) for SP and 96.9% (95% CI 67%–99.8%) for SU, respectively, both with considerable heterogeneity (I² = 77.7% with SP and 85.3% with SU) (Supplementary Fig. 4a and Supplementary Fig. 4b). The pre-planned sensitivity analyses did not explain heterogeneity, nor suggested different results in subgroups.

Adverse events

Only five of the included studies [10, 11, 17, 20, 21] reported data on the rate of adverse events (AEs), being 0% in all but one study, for both techniques. One study [21] reported early adverse events in two of 50 cases (4%) being rigors without fever in one case and mild abdominal pain in the other.

Discussion

EUS is the preferred method to obtain cytopathological characterization of pancreatic solid lesions. The best technique to obtain diagnostic samples is often debated with different possible variables in terms of choice of the needle and of sampling technique [2]. In the past decade a number of studies [3, 22] have demonstrated that the Suction technique is superior to non-Suction in terms of sensitivity accuracy. The SP technique has been subsequently investigated in non-randomized studies, both with FNB (Procore) [23] and FNA [4] needles, with results suggesting high diagnostic yield.

More recently, the SU and SP techniques have been compared in several RCTs. The results of these studies were inconclusive, possibly due to the relatively small number of enrolled patients. The present systematic review and meta-analysis was aimed at gathering data from prospective randomized controlled trials comparing the diagnostic adequacy and accuracy of SU and SP. Indeed, while both these techniques are widely employed, the most recent ESGE guidelines suggested to employ SU as evidence on the use of SP was limited at the time of their preparation.

Our results suggest that SP has similar adequacy and accuracy when compared to SU. Moreover, as expected, SP might result in a lower degree of blood contamination.

We have been able to analyze data from seven studies, two of which were only available as abstracts when this manuscript was submitted, also became available as full-text during the revision process. The studies were heterogeneous in terms of design, with four of them randomizing distinct patients to each technique arm and the remaining three sampling the same lesion with both techniques in a randomized order for a total of 475 patients, with 163 lesions sampled with SU, 164 with SP and 148 by both. Also, the end-points evaluated were different, as some studies had adequacy and other accuracy as primary outcome, thus only a portion of the studies were included in each analysis.

Adequacy of the samples was the first outcome we aimed at evaluating. Only four studies provided data for this analysis, with results suggesting no differences between the two techniques with an OR of 0.98 (95% CI 0.44–2.17) (Fig. 2). Notably, there was no publication bias, nor heterogeneity for this analysis, definition of adequacy was similar and the employed sensitivity analyses also showed no significant differences in adequacy between the two techniques. However, there was a non-significant trend toward a higher adequacy of FNA needles with suction, while this was not the case for FNB needles. Whether this non-significant difference is incidental or it has a technical explanation is difficult to be assessed.

We also evaluated the pooled rate of high blood contamination with data extracted from 5 studies, with results showing a higher rate with SU (27.6%) compared to SP (19.7%) (Supplementary Fig. 1a and Supplementary Fig. 1b). Although the heterogeneous definitions of “blood contamination” in the different studies urge caution, this result is not surprising, as suction, for its higher force of aspiration, is known to cause a higher amount of blood compared to non-suction techniques; in a
recent non-randomized study comparing SU and SP, Bor et al. [9] also reported that the two techniques have similar yield but SP has lower bloodiness of samples. While this does not seem to interfere with the diagnostic adequacy in the studies we analyzed, blood contamination might significantly interfere with macroscopic on-site evaluation (MOSE) which is becoming an increasingly popular tool to increase the quality of EUS-FNB when ROSE is not available [24].

Furthermore, as EUS-guided sampling is becoming important to provide material for molecular tests, the possibility to avoid a significant contamination with blood might turn out to be a critical issue in the future [25].

As far as regards the accuracy of SU and SP, four studies provided sufficient data for a comparison between the two methods (Fig. 3), with results suggesting no differences, with a non-significant trend in favor of SP (OR = 0.82; 95% CI 0.36–1.85; P = 0.63) and moderate heterogeneity that was not explained by the sensitivity analyzes. The single study conducted in Asia with these data available showed a significantly higher accuracy with SP [18]. The pooled rate of diagnostic accuracy with the 2 techniques resulted to be 89.5% with SP and 48.6% with SU. If this small 5% difference in favour of SP is true, setting an alpha error of 0.05 and a beta error of 0.20, it would require a RCT of 982 patients per arm to be confirmed with a statistical significance.

Strengths of the present study are its novelty dealing with a hot and controversial topic, the rigorous methodology, the absence of publication bias and the attempt to take into account several factors that might have caused heterogeneity, including the size of the employed needles. A recent meta-analysis did not demonstrate significant difference between a core needle and standard FNA needle possibly supporting such findings [26].

Although a meta-regression would have been a more reliable method to take into account heterogeneity, unfortunately such analysis is unfeasible due to the low number of studies, as the Cochrane guidelines suggest performing meta-regression with a minimum of 11 studies.

However, several limitations have to be considered. First of all, the relatively low number of included studies and of enrolled patients urge caution in the interpretation of the present results. On the other hand, the choice to limit the analysis to RCTs makes the results more reliable and robust. Second, the included studies were heterogeneous in terms of design, choice and definition of the outcomes and technical aspects. Despite this, the heterogeneity was low in most instances. Also, we have not been able to fully take into account variables such as the importance of ROSE (as it was available in only one study), of fanning and actual speed of needle movements (as it was not always clearly stated) or time of application of suction technique or number of to-and-fro movements. As an example, as far as regards fanning, while one study [18] included fanning in the SP but not in the SU arm, in other studies fanning was either “performed at the discretion of the endoscopist” [10] or performed in both arms “if possible” [17], while no specific information was provided in other studies. Also, other factors as the site, size and histologic nature of the lesions might affect the results. It might, indeed, be possible that one method is superior to the other in a certain setting, but there was no sufficient information to investigate this aspect.

Conclusion

In conclusion, despite these weaknesses, the current study gathers for the first time available evidence from RCTs comparing SU and SP and suggests that sampling of pancreatic solid lesions with the SP technique is not inferior to SU, thus no evidence-based suggestion on the use of one or the other should be given.

Most likely, RCTs aimed at demonstrating the statistical superiority of one of the two methods should be extremely large, multicenter studies with careful methodology. However, such studies might prove to be clinically non-relevant.

Competing interests

The authors declare that they have no conflict of interest.

References


