PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	New Onset of DiabetEs in aSsociation with pancreatic ductal adenocarcinoma (NODES trial): Protocol of a Prospective, Multicentre Observational trial
AUTHORS	Illés, Dóra; Iványi, Emese; Holzinger, Gábor; Kosár, Klára; Gordian, Adam M.; Kamlage, Beate; Zsóri, Gábor; Tajti, Máté; Svébis, Márk; Horváth, Viktor; Oláh, Ilona; Márta, Katalin; Váncsa, Szilárd; Zádori, Noémi; Szentesi, Andrea; Czakó, Bálint; Hegyi, Péter; Czakó, László

VERSION 1 – REVIEW

REVIEWER	Tianpei Hong Department of Endocrinology and Metabolism, Peking University Third Hospital, Beijing, China
REVIEW RETURNED	23-Mar-2020

GENERAL COMMENTS	This manuscript reported the research protocol of NODES, a prospective, multi-centre observational cohort study. The aims of this project are to estimate the incidence of pancreatic ductal adenocarcinoma in patients with new-onset diabetes, and validate a biomarker that distinguishes patients with PaC-caused type 3c diabetes mellitus from patients with type 2 diabetes mellitus. This study would be helpful for clinicians to diagnose pancreatic ductal adenocarcinoma in an early operable stage. Major comments: 1. Sample size: One of the main aims of this study is to validate the biomarker panel that is suitable for early stage diagnosis of PaC. However, only 20 PaC cases are expected during the follow-up time. Please clarify the effectiveness of this trial in regarding to this purpose. 2. Sample size: “Considering 105 drop-out rate, ……”. Do you means 5% drop-out rate? 3. Sample size: Please describe the sample size calculation method in more detail. 4. Control group: “…... 250 matched patients in the control group……”. What is the meaning of this description? Where are the inclusion and exclusion criteria of the control group. 5. Inclusion criteria: Please explain why you choose patients over 60 years of ages as the target population in this study, which is not completely consistent with the risk population (at age 50 or older) as described in the Introduction. 6. The format of references is not consistent. Please unify and modify based on BMJ style. Minor comments:

	7. Page 9, Line 36: “ambigous lesions” should read as “ambiguous lesions”. 8. Page 6, Line 24-26: “……(the value of the Standardised Incidence Ratio for PaC in new-onset type 2 diabetic patients was 198.6 (95% CI = 6.25-46.9)); ……”. Please confirm if the numerical value 198.6 is correct. 9. In Main text and Figure legends, “-80°C” should read as “□80°C”. 10. In the legend of Figure 2, “visit0” should read as “visit 0”. 11. Figure 2, the box of Form B: “clinical symptoms” seem to be “Clinical data” as in the box of Form A. In addition, the first letter should be upper-case as in the box of Form A. 12. Figure 2, the box of Form A and Form B: body height or BMI should be added.

REVIEWER	Dana K. Andersen, M.D. Division of Digestive Diseases and Nutrition National Institute of Diabetes and Digestive and Kidney Diseases National Institutes of Health Bethesda, Maryland USA
REVIEW RETURNED	14-Apr-2020

GENERAL COMMENTS	The manuscript by Illes et al. describes a prospective, observational, multi-center study designed to identify early stage pancreatic ductal adenocarcinoma in subjects with new onset diabetes over age 60 who are followed and studied repeatedly over 3 years. The study, termed the NODES trial, is already ongoing and presumably recruiting subjects. Therefore deficiencies in the design are not likely to be corrected, but further clarification of details for the purpose of publishing this description should be addressed. Introduction: Unfortunately, some of the statistics cited in the introduction are incorrect or badly outdated. The disease of focus is pancreatic ductal adenocarcinoma but the title of the manuscript and the introduction refers only “pancreatic cancer.” It is not until a description of the secondary outcomes that the most specific description of pancreatic ductal adenocarcinoma is used. It would be better to use the specific terminology from the beginning. Introduction: The cited 5-year survival of 6% is less than the 8-9% 5-year survivals being reported currently. This misrepresentation is due to the fact that the source of the 6% figure is a report published in 2013. Surely the authors can do a better job of providing current information. Introduction: To say that the survival has not changed significantly over the past forty years ignores the fact that from 1971, when the 5 year survival of pancreatic ductal adenocarcinoma was 4%, to the present, when the survival is 8-9%, indicates that survival has in fact doubled. Although still regrettably low, this increase is driving the current interest in early detection and should be acknowledged. Introduction: The statement that pancreatic cancer “…is projected to be the third leading cause of cancer-related death by 2030” is incorrect. Pancreatic ductal adenocarcinoma is currently the third leading cause of cancer-related death, lagging behind lung cancer and colorectal cancer, and ahead of breast cancer. It is projected to overtake the rate of death due to colorectal cancer by 2030 and then will be the second leading cause of cancer-related death. Pancreatic cancer and diabetes mellitus: Although the authors acknowledge that the diabetes cause by pancreatic ductal

	adenocarcinoma is actually “diabetes type 3c (T3cDM),” they failed to account for this is their prior study (Pancreatology 2016) where the investigators assumed that all new onset diabetes was Type 2 (T2DM), and never attempted to discriminate T2DM from T3cDM. This is a common problem with most of the published studies of the relationship of pancreatic cancer and diabetes, and it is disappointing that the current NODES trial perpetuates this omission. The work of Ewald and Hardt has shown that the risk of pancreatic ductal adenocarcinoma is nearly 10-fold higher in T3cDM than in T2DM, so this is precisely the form of diabetes that one would wish to identify as the highest-high risk group. The authors should provide some explanation as to why they have not taken this into consideration in the design of their trial. Methods and analysis: Table 1 indicates that the age criterion for entry into the study is 60 years. It is not explained how this age cut off was chosen, or why a lower cutoff of, say, 50 years was not selected. Methods and analysis: Table 1 indicates that a random plasma glucose level (>200mg/dL) will be used to identify diabetes, but diabetologists have noted that random plasma glucose levels are unreliable and a poor criterion precisely because they are random and unreproducible. It would be better to eliminate this criterion or explain how subjects who are identified by this criterion will have their diabetes status verified by some other test. Methods and analysis: Table 1 indicates that an exclusion criterion for subject recruitment is “chronic pancreatitis.” It has been shown in several international studies that the combination of chronic pancreatitis and diabetes defines one of the highest risk categories for the development of pancreatic ductal adenocarcinoma, with relative risk values that are increased from 12- to 33-fold above the general population. Chronic pancreatitis also accounts for the majority of patients with T3cDM, so including this group would likely increase the number of subjects who are found to develop pancreatic adenocarcinoma during the 3 year follow up period of the study. The authors should explain why they have excluded this high risk group from their study. Methods and analysis: The sample size description is completely inadequate. How many subjects are planned to be recruited? How many are expected to develop pancreatic ductal adenocarcinoma? How many are expected to drop out or be lost to follow up? If this is a three year study, how many subjects will be recruited in Year 1? Are all recruitments planned to be completed in Year 1? What happens when a subject is found to develop diabetes in Year 2? Are they then followed for only 1 year, or for 2 more years? Methods and analysis: The study protocol describes plans to measure C-peptide and anti-GAD antibodies, presumably to identify type 1 diabetes (T1DM), but it is unclear whether those subjects will be included in the observational cohort. Also no tests are described to identify T2DM from T3cDM (see discussion above) which seems to be a recurring problem. There have now been several test methods described which offer this ability. The authors should explain why they have ignored this discrimination.

VERSION 1 – AUTHOR RESPONSE

1. **Reviewer:**

Major comments:

1. Sample size: One of the main aims of this study is to validate the biomarker panel that is suitable for early stage diagnosis of PaC. However, only 20 PaC cases are expected during the follow-up time. Please clarify the effectiveness of this trial in regarding to this purpose.

Thank you for your comment. We calculated the sample size with a 10% dropout, 80% power and 95% significance level*. *We extended the sample size calculation part of the manuscript, and we provided the section with relevant reference as a source for our methods.

2. Sample size: "Considering 105 drop-out rate, …...". Do you means 5% drop-out rate?

Thank you for your comment. It was probably just a typo. Now it is corrected: see Methods – Sample size.

3. Sample size: Please describe the sample size calculation method in more detail.

We agree with you. We extended the sample size calculation part of the manuscript, and we provided the section with relevant reference as a source for our methods.

4. Control group: "…... 250 matched patients in the control group…...". What is the meaning of this description? Where are the inclusion and exclusion criteria of the control group.

The inclusion and exclusion criteria of the control group have been added to the Methods – Sample size.

5. Inclusion criteria: Please explain why you choose patients over 60 years of ages as the target population in this study, which is not completely consistent with the risk population (at age 50 or older) as described in the Introduction.

Approximately only 1% of patients with new-onset diabetes at age 50 or older will be diagnosed with PDAC within 3 years of first meeting criteria for diabetes. This low prevalence rate of PDAC would require a great number of new-onset diabetic patients to be screened in our study, which is neither desirable nor cost-effective. To enrich our study population new-onset diabetic patients over 60 years will be selected. The incidence rate of PDAC boosts after the 60 decade (de Souza P. Pancreas 2017;46:699–706, J. Fest et al. European Journal of Cancer 2017;72:86-191). Furthermore, the mean age of PDAC diagnosis with new-onset DM was demonstrated to be 66.9 years (Mizuno S. Pancreas 2014;43: 1014-1017).

Minor comments:

Page 6, Line 24-26: "……(the value of the Standardised Incidence Ratio for PaC in new-onset type 2 diabetic patients was 198.6 (95% CI = 6.25-46.9)); ……". Please confirm if the numerical value 198.6 is correct.

Yes, the value of the Standardised Incidence Ratio (SIR) in our previous study was 198.6 with 95% Confidence interval ranges between 6.25-46.9. SIR was calculated with the person-time incidence based on our study (2%, calculated from the general incidence which was 2.78%=3 PDAC/108 cases) and the age-adjusted incidence of PDAC in Hungary (9.3 cases/100.000 persons) upon the given formula: https://www.state.nj.us/health/ceohs/documents/eohap/haz_sites/passaic/pompton_lakes/pompton_lakes_fs_sir.pdf

2. Reviewer

1. Introduction: Unfortunately, some of the statistics cited in the introduction are incorrect or badly outdated. The disease of focus is pancreatic ductal adenocarcinoma but the title of the manuscript and the introduction refers only "pancreatic cancer." It is not until a description of

the secondary outcomes that the most specific description of pancreatic ductal adenocarcinoma is used. It would be better to use the specific terminology from the beginning.

Thank you for this remark, the terminology is unified as "pancreatic ductal adenocarcinoma (PDAC)".

2. Introduction: The cited 5-year survival of 6% is less than the 8-9% 5-year survivals being reported currently. This misrepresentation is due to the fact that the source of the 6% figure is a report published in 2013. Surely the authors can do a better job of providing current information.
3. Introduction: To say that the survival has not changed significantly over the past forty years ignores the fact that from 1971, when the 5 year survival of pancreatic ductal adenocarcinoma was 4%, to the present, when the survival is 8-9%, indicates that survival has in fact doubled. Although still regrettably low, this increase is driving the current interest in early detection and should be acknowledged.
4. Introduction: The statement that pancreatic cancer "…is projected to be the third leading cause of cancer-related death by 2030" is incorrect. Pancreatic ductal adenocarcinoma is currently the third leading cause of cancer-related death, lagging behind lung cancer and colorectal cancer, and ahead of breast cancer. It is projected to overtake the rate of death due to colorectal cancer by 2030 and then will be the second leading cause of cancer-related death.

The epidemiological data and the relevant literature of PDAC were updated upon the suggestions.

5. Pancreatic cancer and diabetes mellitus: Although the authors acknowledge that the diabetes cause by pancreatic ductal adenocarcinoma is actually "diabetes type 3c (T3cDM)," they failed to account for this is their prior study (Pancreatology 2016) where the investigators assumed that all new onset diabetes was Type 2 (T2DM), and never attempted to discriminate T2DM from T3cDM. This is a common problem with most of the published studies of the relationship of pancreatic cancer and diabetes, and it is disappointing that the current NODES trial perpetuates this omission. The work of Ewald and Hardt has shown that the risk of pancreatic ductal adenocarcinoma is nearly 10-fold higher in T3cDM than in T2DM, so this is precisely the form of diabetes that one would wish to identify as the highest-high risk group. The authors should provide some explanation as to why they have not taken this into consideration in the design of their trial.

This point correlates to the last one, therefore the answers to these two questions will be interpreted together.

6. Methods and analysis: Table 1 indicates that the age criterion for entry into the study is 60 years. It is not explained how this age cut off was chosen, or why a lower cutoff of, say, 50 years was not selected.

Approximately only 1% of patients with new-onset diabetes at age 50 or older will be diagnosed with PDAC within 3 years of first meeting criteria for diabetes. This low prevalence rate of PDAC would require a great number of new-onset diabetic patients to be screened in our study, which is neither desirable nor cost-effective. To enrich our study population new-onset diabetic patients over 60 years will be selected. The incidence rate of PDAC boosts after the 60 decade (de Souza P. Pancreas 2017;46:699–706, J. Fest et al. European Journal of Cancer 2017;72:86-191). Furthermore, the mean age of PDAC diagnosis with new-onset DM was demonstrated to be 66.9 years (Mizuno S. Pancreas 2014;43: 1014-1017).

7. Methods and analysis: Table 1 indicates that a random plasma glucose level (>200mg/dL) will be used to identify diabetes, but diabetologists have noted that random plasma glucose levels are unreliable and a poor criterion precisely because they are random and unreproducible. It

would be better to eliminate this criterion or explain how subjects who are identified by this criterion will have their diabetes status verified by some other test.

Thank you for this remark. The random plasma glucose value is not used for diagnosing diabetes in Hungary. We indicated all the diabetes-diagnosing criteria suggested by the American Diabetes for the sake of completeness, however, measuring random plasma glucose value does not belong to the accepted diabetes-diagnosing methods of the study. Therefore, random plasma glucose was eliminated from Table 1.

8. Methods and analysis: Table 1 indicates that an exclusion criterion for subject recruitment is "chronic pancreatitis." It has been shown in several international studies that the combination of chronic pancreatitis and diabetes defines one of the highest risk categories for the development of pancreatic ductal adenocarcinoma, with relative risk values that are increased from 12- to 33-fold above the general population. Chronic pancreatitis also accounts for the majority of patients with T3cDM, so including this group would likely increase the number of subjects who are found to develop pancreatic adenocarcinoma during the 3 year follow up period of the study. The authors should explain why they have excluded this high risk group from their study.

One of the main purposes of this study is to validate a biomarker panel for whether it is able to differentiate PDAC-caused diabetes (PDAC-T3cDM) from "simple" type-2 diabetes mellitus. The biomarker panel used in this study is already validated for distinguishing between PDAC and chronic pancreatitis (CP) (Mayerle J. et al. Identification of plasma metabolites as biomarker candidates for the diagnosis of pancreatic ductal adenocarcinoma).It is true that including CP patients could multiply the number of cases diagnosed with PDAC in the study, but it also would lead to misinterpretation of the data.

9. Methods and analysis: The sample size description is completely inadequate. How many subjects are planned to be recruited? How many are expected to develop pancreatic ductal adenocarcinoma? How many are expected to drop out or be lost to follow up? If this is a three year study, how many subjects will be recruited in Year 1? Are all recruitments planned to be completed in Year 1? What happens when a subject is found to develop diabetes in Year 2? Are they then followed for only 1 year, or for 2 more years?

Yes, you are right. As mentioned in our previous answers, we extended the sample size calculation, and we provided the section with relevant reference as a source for our methods.

10. Methods and analysis: The study protocol describes plans to measure C-peptide and anti-GAD antibodies, presumably to identify type 1 diabetes (T1DM), but it is unclear whether those subjects will be included in the observational cohort. Also no tests are described to identify T2DM from T3cDM (see discussion above) which seems to be a recurring problem. There have now been several test methods described which offer this ability. The authors should explain why they have ignored this discrimination.

Evaluating C-peptide and anti-GAD antibodies assist us for differentiating type-1 diabetes mellitus (T1DM). Patients with T1DM will be excluded from the study, therefore T1DM was added as exclusion criterion to the "Exclusion criteria" in the Methods section. It is true that T3cDM means the highest-high risk group for PDAC and it is still underdiagnosed in the clinical practice –maybe because its diagnosis is based on complex, expensive tests that are not routinely available (Ewald N, Bretzel RG. Diabetes mellitus secondary to pancreatic diseases (type 3c)–are we neglecting an important disease? Eur J Intern Med 2013;24:203–206). To diagnose T3cDM patients based on the criteria mentioned above would lead to enormous difficulties in inclusion of subjects and it would not be a cost-effective screening method, which is unfavorable. Beside, not every patients with PDAC-

T3cDM fulfill the criteria of T3cDM. One of the main purposes of this study is to validate a biomarker panel for whether it is able to differentiate PDAC-caused diabetes (PDAC-T3cDM) from "simple" new-onset, not-type1 diabetes mellitus. In that manner, this biomarker panel could be a diagnostic tool for the T3cDM-subgroup PDAC-T3cDM.

VERSION 2 – REVIEW

REVIEWER	Tianpei Hong Peking University Third Hospital, China
REVIEW RETURNED	31-May-2020

GENERAL COMMENTS	Major comments: 1. Sample size: The authors have modified a lot to make the calculation more clearly. However, since the meaning of the "case" and "control" groups is not defined clearly in this manuscript, it is hard to understand how the ratio is calculated in the sentence "……in the case and the control groups (Case: PDAC 1%, non-PDAC: 99%; Control: PDAC 0.125%, non-PDAC 99.875%).". It is also very confused whether the "control" group in the sentence "250 patients with non-pancreatic, non-malignant gastrointestinal diseases without diabetes serve as control group" has the same meaning as above. The definition of the groups should be clearly described. 2. Sample size: As my understanding in this part, the calculated sample size in case group is 2552, while the sample size in control group is 250, and the ratio is 10:1 for the case group over control group. The ratio is uncommon in most of the studies, and should be explained clearly. Meanwhile, the 1:1 here in the manuscript seems to be a typo and should be modified or explained. 3. Sample size: The inclusion and exclusion criteria of the control group should be described in more detail as it is still not clear enough to be mentioned in a simple sentence. Whether the inclusion and exclusion criteria of the case group is also applicable for the control group should be described. Minor comments: 1. Page 6, Line 25-28: "……(the value of the Standardised Incidence Ratio for PDAC in new-onset type 2 diabetic patients was 198.6 (95% CI = 6.25-46.9)); ……". Although the authors have explained how the standardised incidence ratio (SIR) is achieved, the SIR should be in the range of 95% CI. Moreover, the calculated SIR should be 298.7 using the number provided by the authors. It seems that the SIR and 95% CI are incorrectly calculated. 2. Page 8, Line 40-42: "250 patients with …… serve as control group." should be read as "Two hundred fifty patients with …… serve as control group.".

REVIEWER	Dana K. Andersen, M.D. Division of Digestive Diseases and Nutrition National Institute of Diabetes and Digestive and Kidney Diseases National Institutes of Health United States
REVIEW RETURNED	18-May-2020

GENERAL COMMENTS	This revised manuscript by Illes et al. describes the multicenter observational cohort study "New Onset of DiabetEs in aSsociation with pancreatic ductal adenocarcinoma (NODES) trial. The manuscript has been improved by attention to issues raised in the first line of review. Some minor flaws persist, however. P6, lines 10-12: "…several diabetes in PDAC patients…" needs revision. P6, line 24: "In our recent prospective study, the prevalence of PDAC in patients with new-onset type 2 diabetes (T2DM) was significantly higher…" implies that only T2DM patients were assessed. But the study referenced did not attempt to discriminate T2DM from T3cDM. Just a few lines before, the authors state "…diabetes in PDAC patients is caused by the cancer itself. In these cases, patients are actually suffering from diabetes type 3c (T3cDM)." To resolve this inconsistency I would suggest revising line 24 to say "In our recent prospective study, the prevalence of PDAC in patients with new-onset diabetes was significantly higher…" This would be factually correct and would avoid the issue of distinguishing T2DM from T3cDM. P13, line 23: In Discussion, the authors should explain why there is no effort made to discriminate T3cDM from T2DM (several criteria for the diagnosis of T3cDM have been proposed), as T3cDM would appear to be the form of diabetes that they are actually searching for. In addition, the authors should explain why they exclude patients with a known history of chronic pancreatitis, when chronic pancreatitis accompanied by diabetes has been shown by several authors to be a very high risk category for PDAC.

VERSION 2 – AUTHOR RESPONSE

3. Reviewer:

Major comments:

1. Sample size: The authors have modified a lot to make the calculation more clearly. However, since the meaning of the "case" and "control" groups is not defined clearly in this manuscript, it is hard to understand how the ratio is calculated in the sentence "……in the case and the control groups (Case: PDAC 1%, non-PDAC: 99%; Control: PDAC 0.125%, non-PDAC 99.875%).". It is also very confused whether the "control" group in the sentence "250 patients with non-pancreatic, non-malignant gastrointestinal diseases without diabetes serve as control group" has the same meaning as above. The definition of the groups should be clearly described.

Thank you for alerting us that the description of the studied population is confusing, you are right. We plan to enroll elderly patients with new onset diabetes, the case and control groups will be the ones with PDAC and those without. We do not plan to include other control patients, which means that the inclusion and exclusion criteria for the "case" and the "control" groups are the same.

2. Sample size: As my understanding in this part, the calculated sample size in case group is 2552, while the sample size in control group is 250, and the ratio is 10:1 for the case group over control group. The ratio is uncommon in most of the studies, and should be explained clearly. Meanwhile, the 1:1 here in the manuscript seems to be a typo and should be modified or explained.

Thank you for your comment on this section. We reconsidered the sample size calculation and provided a different sample size based on the hypothesis on sensitivity and specificity of the biomarker signature. Based on previous data and on an assumed prevalence of PDAC in the given population, we calculated a sample size of 1241 that needs to be enrolled to reach statistical significance. Now it is explained in the Sample size section.

3. Sample size: The inclusion and exclusion criteria of the control group should be described in more detail as it is still not clear enough to be mentioned in a simple sentence. Whether the inclusion and exclusion criteria of the case group is also applicable for the control group should be described.

Thank you for your comment. This question is answered together with the first one.

Minor comments:

1. Page 6, Line 25-28: "……(the value of the Standardised Incidence Ratio for PDAC in new-onset type 2 diabetic patients was 198.6 (95% CI = 6.25-46.9)); ……". Although the authors have explained how the standardised incidence ratio (SIR) is achieved, the SIR should be in the range of 95% CI. Moreover, the calculated SIR should be 298.7 using the number provided by the authors. It seems that the SIR and 95% CI are incorrectly calculated.

Thank you for this comment. You are right, the exact method for calculating SIR was not clearly detailed in our previous study. We calculated SIR based on person-time incidences as follow: the cumulative follow-up time was 162.42 years for the 108 included patients. The calculation of the cumulative follow – up period was made upon this example https://www.ctspedia.org/do/view/CTSpedia/StudyAnalysisPerson (see Method I.). 3 patients had PDAC in our study, so the person-time incidence is 3/162.42*100 = 1.847 cases per 100 personyears.
The age-adjusted incidence for PDAC in Hungary was 9.3/ 100000 personyears, which is equal to 0.0093/ 100 personyears. Using this data SIR in our study is 1.847/0.0093= 198.6.

About the correlation of Confidence Interval and SIR (Source: Massachusetts Department of Public Health, Bureau of Environmental Health Assessment (December 1998)):

"An SIR greater than 100 indicates that more cancer cases occurred than expected and an SIR less than 100 indicates that fewer cancer cases occurred than expected. To determine if the observed number of cases is significantly different from the expected number or if the difference may be due solely to chance, a 95% confidence interval (CI) was calculated for each SIR. Specifically, a 95% CI is the range of estimated SIR values that has a 95% probability of including the true SIR for the population. If the 95% CI range does not include the value 100, then the study population is significantly different from the comparison or "normal" population. "Significantly different" means there is less than 5% percent chance that the observed difference is the result of random fluctuation in the number of observed cancer cases."

Which means that the studied subgroup "patients with new-onset diabetes" has a significant higher risk for developing PDAC, so they count as a high risk group worth for screening for PDAC.

2. Reviewer:

Thank you for your recommendations we are fully agreed with. We made corrections upon your suggestions.

More detailed, in response regarding the last two remarks: Thank you for alerting us that we did not make ourselves clear in this part of the discussion. The NODES study does not aim to discriminate all T3cDM from T2DM, as PDCA is only one of several potential causes of T3cDM. One of the aims of the study is to estimate the incidence of PDAC in new-onset of diabetes and to evaluate the diagnostic potential of a biomarker panel in the population of new-onset of diabetes.
The interaction between chronic pancreatitis, pancreatic cancer, and diabetes are complex and not fully understood (Andersen DK, Korc M, Petersen GM, et al. Diabetes, Pancreatogenic Diabetes, and Pancreatic Cancer. _Diabetes_. 2017;66(5):1103-1110. doi:10.2337/db16-1477). A high number of chronic pancreatitis develop diabetes mellitus type 3c in the course of the disease. Though several risk factors of pancreatic cancer are known, among others chronic pancreatitis, most of the pancreatic cancer cases are sporadic and new-onset of diabetes mellitus can be a first symptom of a pancreatic cancer. In these cases, the diabetes mellitus is also referred to a type3c, however, it is not known whether the molecular pathomechanism are different or similar. As chronic pancreatitis has a marked influence on the metabolism (see Lusczek ER, Paulo JA, Saltzman JR, Kadiyala V, Banks PA, Beilman G, et al. Urinary 1H-NMR metabolomics can distinguish pancreatitis patients from healthy controls. Jop 2013;14:161-70, and Ouyang D. Metabolomic characterization of human pancreatitis by ^{1}H-NMR spectroscopy. Hepatogastroenterology 2012;59:2314-7), the inclusion of CP patients could lead to a bias, entailing the risk that any differentiating metabolite signature would actually differentiate between healthy diabetics and CP patients with PDAC. This, however, is not the aim of the NODES study. The metabolic differences between CP and PDAC have been investigated and a metabolic signature for differentiating between these diseases has been published (Mayerle J, Kalthoff H, Reszka R, et al. Metabolic biomarker signature to differentiate pancreatic ductal adenocarcinoma from chronic pancreatitis. Gut 2018;67(1):128-137.). Following up on this study, the metabolic differences between CP and PDAC patients are being investigated by the META-PAC consortium in Germany, a large study including more than 1000 patients that has been running since 2015 and will be completed 2022. The NODES study has a different approach, focusing on the metabolic differences between PDAC-caused diabetes and other new-onset diabetes in the elderly population. A metabolic signature differentiating between these groups could be used to develop a screening test for early identification of PDAC patients in the risk group elderly new onset of diabetes patients. Besides, since chronic pancreatitis patients have other incidences of PDAC, inclusion of chronic pancreatitis patients would result in the need to be separately addressed in the biostatistical analysis and this would increase the number of subjects to be included. New onset of diabetes subjects at advanced age usually show up in primary care. Hereditary chronic pancreatitis patients experience the onset of the disease in childhood and are usually in regular medical care of gastroenterologists for their entire life. They have the highest risk among chronic pancreatis subjects for PDAC and do not show up with new-onset of diabetes in the primary care.Ethanol abuse-induced chronic pancreatitis patients usually show up in emergency rooms for pain release. That means that the clinical use cases are separated and the usefullness of the biomarker has to to be evaluated separately.
A sentence to clarify this has now been added to the discussion section of the manuscript.

VERSION 3 – REVIEW

REVIEWER	Tianpei Hong Director, Department of Endocrinology & Metabolism, Peking University Third Hospital, Beijing, China
REVIEW RETURNED	05-Aug-2020

GENERAL COMMENTS	The revised manuscript has been improved by attention to issues raised by the reviewers. The authors have provided a different sample size calculation based on the hypothesis on sensitivity and specificity of the biomarker signature. However, the prevalence of PDAC in the given population assumed by the authors is 4%, which seems too high even for this given population with age over 60, with diabetes and in Hungary. In reference 9, Chari et al. reported a prevalence of 0.85% for pancreatic cancer within 3 years of meeting criteria for diabetes with age over 50 in the USA. In reference 29, the age-standardized rate incidence for pancreatic cancer was 7.6 per 100,000 people in North America and 10.8 per 100,000 people in Hungary. The authors should explain how the prevalence of 4% for PDAC is assumed.

REVIEWER	Dana K. Andersen, M.D. NIDDK/NIH USA
REVIEW RETURNED	31-Jul-2020

GENERAL COMMENTS	The revised manuscript by Illes et al. adequately addresses the concerns and recommendations of prior reviews. The one potential pitfall of the study, not addressed by the authors, is the plan to assess the metabolic biomarker panel only once at entry into the study. If a subject were to have a non-diagnostic or negative result on the first visit, but develop a diagnostic or abnormal result 1 or 2 years later, this would be missed.

VERSION 3 – AUTHOR RESPONSE

1. **Reviewer:**

Major comment:

> The revised manuscript has been improved by attention to issues raised by the reviewers. The authors have provided a different sample size calculation based on the hypothesis on sensitivity and specificity of the biomarker signature. However, the prevalence of PDAC in the given population assumed by the authors is 4%, which seems too high even for this given population with age over 60, with diabetes and in Hungary. In reference 9, Chari et al. reported a prevalence of 0.85% for pancreatic cancer within 3 years of meeting criteria for diabetes with age over 50 in the USA. In reference 29, the age-standardized rate incidence for pancreatic cancer was 7.6 per 100,000 people in North America and 10.8 per 100,000 people in Hungary. The authors should explain how the prevalence of 4% for PDAC is assumed.

Thank you for alerting us that the value of the estimated prevalence seems too high. You are absolutely right, the 4% was a result of a communication gap between us and the statistician team. Unfortunately, there is almost no data available regarding the rate of pancreatic cancer in patients diagnosed with diabetes mellitus in Hungary. Based on the data of international studies, the prevalence of pancreatic cancer in diabetic patients ranges 1-2% (Magruder JT et al. Pancreas 2011 Apr;40(3):339-51.; Hong SG et al. Korean J Gastroenterol 2009 Sep;54(3):167-73). In our previous study, the prevalence of pancreatic cancer was 2.78% among patients with newly diagnosed diabetes treated at our clinic (Illés D et al. Pancreatology Mar-Apr 2016;16(2):266-71). Considering these data and the high incidence rate in Hungary, we re-evaluated the case-number calculation based on an estimated 2% prevalence.

2. **Reviewer:**

The revised manuscript by Illes et al. adequately addresses the concerns and recommendations of prior reviews. The one potential pitfall of the study, not addressed by the authors, is the plan to assess the metabolic biomarker panel only once at entry into the study. If a subject were to have a non-diagnostic or negative result on the first visit, but develop a diagnostic or abnormal result 1 or 2 years later, this would be missed.

Thank you for your comment, you're absolutely right: that is why we are planning to send blood sample to biobank for assessing the biomarker panel in every 12 months, as it is written in the "Study protocol" section and marked on Figure 1.

VERSION 4 – REVIEW

REVIEWER	Tianpei Hong Peking University Third Hospital, Beijing, China
REVIEW RETURNED	16-Sep-2020

GENERAL COMMENTS	The revised manuscript adequately addresses the concerns of prior reviews.