

# SCIENTIFIC REPORTS



OPEN

## Combined use of CEMIP and CA 19-9 enhances diagnostic accuracy for pancreatic cancer

Hee Seung Lee<sup>1</sup>, Chan Young Jang<sup>1</sup>, Sun A Kim<sup>1</sup>, Soo Been Park<sup>1</sup>, Dawoon E. Jung<sup>2</sup>, Bo Ok Kim<sup>3</sup>, Ha Yan Kim<sup>3</sup>, Moon Jae Chung<sup>1</sup>, Jeong Youp Park<sup>1</sup>, Seungmin Bang<sup>1</sup>, Seung Woo Park<sup>1</sup> & Si Young Song<sup>1</sup>

Carbohydrate antigen (CA) 19-9 is the only diagnostic marker used in pancreatic cancer despite its limitations. Here, we aimed to identify the diagnostic role of CEMIP (also called KIAA1199) combined with CA 19-9 in patients with pancreatic cancer. A retrospective analysis of prospectively collected patient samples was performed to determine the benefit of diagnostic markers in the diagnosis of pancreatic cancer. We investigated CEMIP and CA 19-9 levels in 324 patients with pancreatic cancer and 49 normal controls using serum enzyme-linked immunosorbent assay. Median CA 19-9 and CEMIP levels were 410.5 U/ml (40.8–3342.5) and 0.67 ng/ml (0.40–1.08), respectively, in patients with pancreatic cancer. The AUROC for CA 19-9 and CEMIP were 0.847 (95% confidence interval [CI]: 0.806–0.888) and 0.760 (95% CI: 0.689–0.831), respectively. Combination of CA 19-9 with CEMIP showed markedly improved AUROC over CA 19-9 alone in pancreatic cancer diagnosis (0.94 vs. 0.89;  $P < 0.0001$ ). CEMIP showed a diagnostic yield of 86.1% (68/79) in CA 19-9 negative pancreatic cancer. Combined use with CEMIP showed significantly improved diagnostic value compared with CA 19-9 alone in pancreatic cancer. Especially, CEMIP may be a complementary marker in pancreatic cancer patients with normal CA 19-9 levels.

Pancreatic cancer is a highly malignant tumor with a poor prognosis. Most patients (80%) present with inoperable advanced pancreatic cancer at the time of initial diagnosis<sup>1</sup>. Although carbohydrate antigen (CA) 19-9 is known as a pancreatic cancer biomarker, it is not commonly used for general screening, owing to its low sensitivity and specificity<sup>2–4</sup>. In particular, false-negative results in the segment of the population with Lewis blood type A-B- and false-positive results in patients with obstructive jaundice limit the specificity of CA 19-9 for pancreatic cancer<sup>5,6</sup>. Therefore, development of novel diagnostic markers is required for the early detection of pancreatic cancer. Previous studies have suggested alternative biomarkers to compensate for the limitations of CA 19-9 in patients with pancreatic cancer<sup>7–9</sup>.

CEMIP was originally identified as a hearing loss-related gene. Increased expression of cell migration-inducing hyaluronan binding protein (CEMIP), a newly identified protein involved in hyaluronan degradation<sup>10</sup>, has been reported in various cancers. CEMIP contributes to breast cancer cell migration with induction of epithelial-mesenchymal transition via calcium signaling<sup>11,12</sup>. Additionally, CEMIP expression is associated with gastric and colorectal tumorigenesis and it is a potential biomarker of gastric cancer<sup>13–15</sup>.

Recently, several studies have investigated the role of CEMIP in pancreatic cancer; CEMIP has been reported to be associated with early detection, cancer cell migration, invasion, and poor prognosis in previous studies<sup>10,16,17</sup>. Suh *et al.* suggested that CEMIP may be useful for detecting pancreatic cancer at an early stage<sup>16</sup>, while Koga *et al.* demonstrated its association with prognosis in pancreatic cancer<sup>10</sup>. However, the sample sizes of these previous studies were too small to draw a solid conclusion<sup>10,16,17</sup>. On the other hand, despite its limitations, CA 19-9 has been proposed as a diagnostic tool for the detection of pancreatic cancer. To date, there have been no studies investigating the diagnostic accuracy of serum CA 19-9 in combination with CEMIP for pancreatic cancer, as compared with serum CA 19-9 alone.

<sup>1</sup>Division of Gastroenterology, Department of Internal Medicine, Yonsei University College of Medicine, Seoul, Korea.

<sup>2</sup>Institute of Gastroenterology, Yonsei University College of Medicine, Seoul, Korea. <sup>3</sup>Biostatistics Collaboration Unit, Yonsei University College of Medicine, Seoul, Korea. Correspondence and requests for materials should be addressed to S.Y.S. (email: [syong@yuhs.ac](mailto:syong@yuhs.ac))

Therefore, we aimed to investigate the differences in CEMIP expression in whole blood between patients with pancreatic cancer and healthy participants, and to identify the role of CEMIP compared with that of CA 19-9 in the diagnosis of pancreatic cancer.

## Methods

**Patients.** We performed a retrospective analysis of prospectively collected patient samples including pancreatic cancer patients and normal individuals between 2007 and 2015. As a control group, normal individuals were defined as people who were not diagnosed with pancreatic cancer, and these individuals were divided into two groups: healthy individuals and patients with benign diseases. Blood samples were obtained immediately following diagnosis and prior to the administration of any oncological or surgical treatment. Information regarding patient demographics and clinical data was obtained from the electronic medical records, including age at diagnosis, sex, location of cancer, serum levels of CA 19-9, antitumor treatment, and tumor stage. Tumor stages were based on the staging classification of the 7th edition of the American Joint Committee on Cancer (AJCC)<sup>18</sup>. Written informed consent was obtained from all patients who approved the use of their blood, and this study was approved by the institutional review board of Yonsei University.

**CA 19-9 and CEMIP.** All methods were performed in accordance with relevant guidelines and regulations. CA 19-9 was measured using chemiluminescence immunoassay on the VITROS 3600 Immunodiagnostic System (Ortho-Clinical Diagnostics Inc., Raritan, NJ, USA). The standard diagnostic cutoff value for CA 19-9 was 37 U/mL. The levels of CEMIP were measured using the enzyme-linked immunosorbent assay (ELISA) method. We detected secreted CEMIP protein in blood samples obtained from patients with pancreatic cancer and individuals without cancer. A 96-well CEMIP ELISA kit (SER965Hu) was purchased from USCN Life Science, Inc. (Wuhan, China) and the procedure was performed according to the manufacturer's instructions. The detection range for the ELISA kit used was 0.156–10 ng/ml. The standard curve concentrations used for the ELISAs were 10, 5, 2.5, 1.25, 0.625, 0.312, and 0.156 ng/ml. The sensitivity of this assay, or lower limit of detection (LLD), was defined as the lowest protein concentration that could be differentiated from zero. It was determined by adding two standard deviations to the mean optical density value of 20 standard replicates and calculating the corresponding concentration. No significant cross-reactivity or interference between CEMIP and analogues was observed.

**Statistical analyses.** All statistical analyses were performed using SPSS for Windows version 20.0 (SPSS, Inc., Chicago, IL, USA), SAS (version 9.4, SAS Inc., Cary, NC, USA), and R package, version 3.0.1. For comparison of serum values between normal individuals and patients with pancreatic cancer, Student's *t* test was employed. The receiver operating characteristic (ROC) curve and the area under the ROC curve (AUROC) were calculated to identify the optimal cut-off CEMIP level associated with diagnosis. We used MedCalc version 11.1 for the ROC analysis (MedCalc Software, Mariakerke, Belgium). The bootstrapping method (1,000 replications) was applied for internal validation of the AUROC. We used Contal and O'Quigley's method of categorizing patients into high or low risk groups for overall survival and chose a cut-off point of CEMIP according to the method<sup>19</sup>. Survival curves were estimated using the Kaplan-Meier method, and the differences among curves were assessed using log-rank tests. A *P*-value < 0.05 was considered to be statistically significant. Univariate and multivariate analysis was performed using Cox regression analysis for risk factors influencing survival in patients who were diagnosed with pancreatic cancer. GraphPad Prism 4 software (GraphPad Software, San Diego, CA, USA) was used for construction of graphs.

## Results

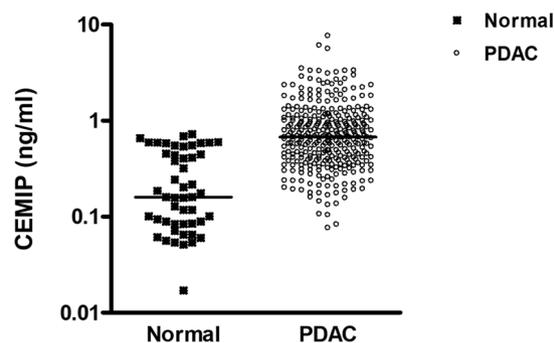
**Patient characteristics.** Serum samples were obtained from 324 patients diagnosed with pancreatic cancer and 49 normal individuals at the Severance Hospital. Normal individuals were divided into two groups: 30 healthy individuals and 19 patients with benign diseases such as common bile duct stone (*n* = 2), gallbladder stone (*n* = 4), intrahepatic stone (*n* = 1), cholecystitis (*n* = 4), liver abscess (*n* = 1), acute hepatitis (*n* = 1), acute pancreatitis (*n* = 1), duodenal polyp (*n* = 1), gastric ulcer (*n* = 1), and benign pancreatic cyst (*n* = 3). Table 1 shows the baseline characteristic of patients in the present study. The median age of patients was 63 years (men, 64.5%). About 75% of patients received palliative chemotherapy, and 170 patients (52%) were diagnosed with stage IV pancreatic cancer. The median overall survival was 314 days (169–581). The median levels of CA 19-9 and CEMIP were 410.5 U/ml (40.8–3342.5) and 0.67 ng/ml (0.40–1.08), respectively, in patients with pancreatic cancer.

**CEMIP as diagnostic marker.** Using the Mann-Whitney *U* test, the CA 19-9 levels were found to be significantly elevated in the serum of patients with pancreatic cancer compared with normal individuals (median, 410.5 vs. 10.8 U/ml, *P* < 0.001). Moreover, CEMIP also exhibited significantly higher expression levels in patients with pancreatic cancer than in normal individuals (0.67 vs. 0.16 ng/ml, *P* < 0.001) (Fig. 1).

The ROC curve analysis indicated the potential diagnostic values of these markers (Fig. 2): The AUROCs for CA 19-9 and CEMIP were 0.85 (95% confidence interval [CI]: 0.806–0.888) and 0.76 (95% CI: 0.689–0.831), respectively. CEMIP also showed good diagnostic performance in internal validation (bootstrap-corrected AUROC, 0.76; 95% CI, 0.692–0.833). Combined ROC curve analysis using CA 19-9 and CEMIP revealed an AUROC of 0.89 (95% CI: 0.862–0.920) with a sensitivity of 80.3% and a specificity of 97.9% in discriminating pancreatic cancer patients from normal individuals. Combination of CA 19-9 with CEMIP showed markedly improved AUROC over that of CA 19-9 alone in the diagnosis of pancreatic cancer against normal individuals (AUROC, 0.89 vs. 0.85; *P* = 0.0119) (Fig. 2A). With combined CEMIP and CA 19-9, the test showed better sensitivity and negative predictive value compared to CA 19-9 alone (Table 2).

| Variables             | Patients (n = 324)  |
|-----------------------|---------------------|
| Age                   | 63.0 ± 10.6         |
| Men                   | 209 (64.5%)         |
| Location              |                     |
| Head                  | 168 (51.9%)         |
| Body                  | 66 (20.4%)          |
| Tail                  | 58 (17.9%)          |
| Mixed                 | 32 (9.8%)           |
| Antitumor treatment   |                     |
| Operation             |                     |
| Whipple procedure     | 4 (1.2%)            |
| PPPD                  | 47 (14.5%)          |
| Total pancreatectomy  | 1 (0.3%)            |
| Distal pancreatectomy | 13 (4.0%)           |
| Chemotherapy          | 245 (75.6%)         |
| Supportive care       | 14 (4.3%)           |
| Stage                 |                     |
| I                     | 4 (1.2%)            |
| II                    | 89 (27.5%)          |
| III                   | 61 (18.8%)          |
| IV                    | 170 (52.5%)         |
| CEMIP, ng/mL          | 0.67 (0.40–1.08)    |
| CA 19-9, U/mL         | 410.5 (40.8–3342.5) |
| Overall survival      | 314 days (169–581)  |

**Table 1.** Baseline characteristics. Variables are expressed as mean ± standard deviation, median (interquartile range), or n (%). PPPD, pylorus-preserving pancreaticoduodenectomy; CEMIP, cell migration-inducing hyaluronan binding protein; CA 19-9, carbohydrate antigen 19-9.



**Figure 1.** Summary of serum CEMIP expression levels detected from pancreatic cancer patients and individuals without cancer. The CEMIP expression was significantly higher in patients with pancreatic cancer than in individuals without cancer ( $P < 0.001$  by paired t-test). The horizontal lines represent the median values. CEMIP, cell migration-inducing hyaluronan binding protein; PDAC, pancreatic ductal adenocarcinoma.

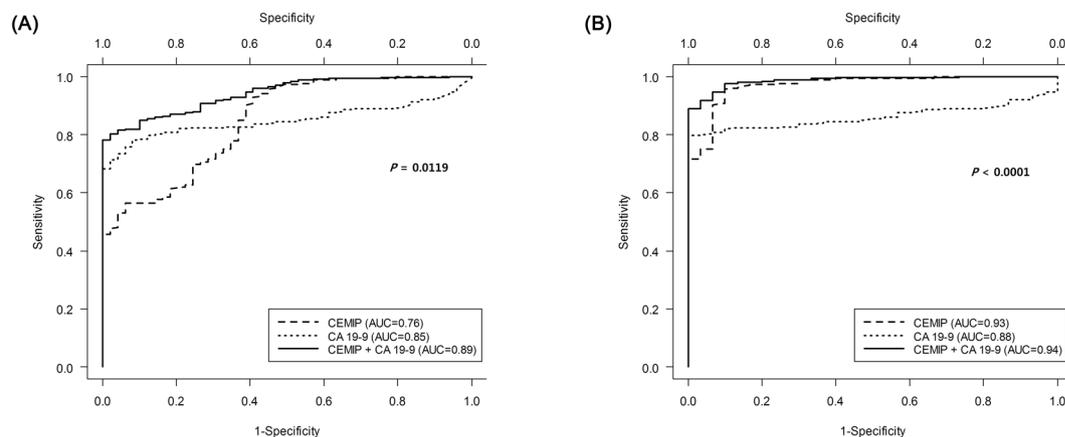
Additional subgroup analysis was performed by including only healthy individuals among all normal individuals. Combination of CA 19-9 with CEMIP showed markedly improved AUROC over that of CA 19-9 alone in the diagnosis of pancreatic cancer against healthy individuals (AUROC, 0.94 vs. 0.88;  $P < 0.0001$ ) (Fig. 2B).

**CEMIP in CA 19-9 negative patients.** Among 324 patients who were diagnosed with pancreatic cancer, 79 patients demonstrated CA 19-9 levels below 37 U/ml, and 19 patients were Lewis blood type A-B-. The proportion of patients who showed high CEMIP levels ( $\geq 0.218$  ng/ml) among pancreatic cancer patients with CA 19-9 in the normal range ( $< 37$  U/ml) was 86.1% (Supplementary Table 1). Among 79 patients with CA 19-9 in the normal range, 68 patients (black colored circle) were diagnosed using CEMIP (diagnostic yield, 86.1%) (Fig. 3).

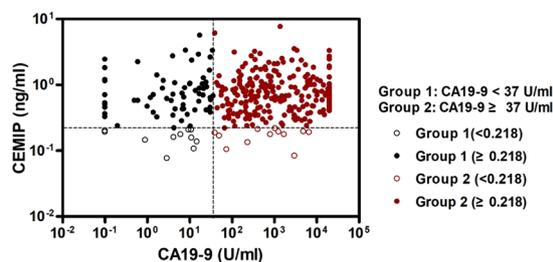
**CEMIP as prognostic predictive marker.** We selected the optimal cut-off value of CEMIP to maximize the survival difference using Contal and O'Quigley's method (CEMIP = 0.429 ng/ml). Using Kaplan-Meier curves, we examined the correlation between CEMIP expression and survival in patients with pancreatic cancer. Survival analysis revealed a significantly shorter overall survival in the high CEMIP expression group ( $\geq 0.429$  ng/ml) compared with the low CEMIP expression group ( $< 0.429$  ng/ml) (9.8 vs. 13.7 months, log-rank test  $P = 0.0175$ ) (Fig. 4).

| Test            | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) |
|-----------------|-----------------|-----------------|---------|---------|
| CA 19-9         | 75.6            | 93.8            | 98.7    | 36.8    |
| CEMIP           | 92.9            | 59.2            | 93.7    | 55.7    |
| CEMIP + CA 19-9 | 96.6            | 59.2            | 94.0    | 72.5    |

**Table 2.** Performance characteristics of CA 19-9 and CEMIP. PPV, Positive predictive value; NPV, Negative predictive value; CA 19-9, Carbohydrate antigen 19-9.

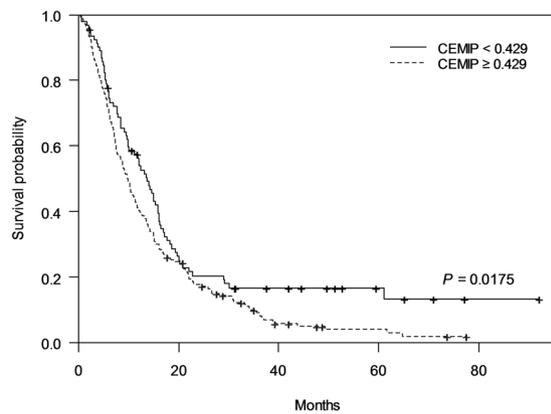


**Figure 2.** ROC curves of CEMIP, CA 19-9, and both to diagnose pancreatic cancer. Normal individuals were defined as people who were not diagnosed with pancreatic cancer and were divided into two groups (healthy individuals and patients with benign disease). (A) Normal individuals including healthy individuals and patients with benign disease were the control group. Combination with CEMIP showed markedly improved AUROC over that of CA 19-9 alone in the diagnosis of pancreatic cancer against normal individuals (AUROC, 0.89 vs. 0.85;  $P = 0.0119$ ). (B) Healthy individuals were the control group. Combination with CEMIP showed markedly improved AUROC over that of CA 19-9 alone in the diagnosis of pancreatic cancer against healthy individuals (AUROC, 0.94 vs. 0.89;  $P < 0.0001$ ). CEMIP, cell migration-inducing hyaluronan binding protein; CA 19-9, carbohydrate antigen 19-9; ROC, receiver operating characteristic; AUC, area under the ROC curve.



**Figure 3.** Distribution of CEMIP and CA 19-9 levels in patients diagnosed with pancreatic cancer. The optimal cut-off levels of CA 19-9 and CEMIP were 37 U/mL and 0.218 ng/mL, respectively. A total of 68 CA 19-9 negative patients (black colored circle) were diagnosed using CEMIP (diagnostic yield, 86.1%). CEMIP, cell migration-inducing hyaluronan binding protein; CA 19-9, carbohydrate antigen 19-9.

To further determine the prognostic value of CEMIP expression for pancreatic cancer, clinicopathological variables as well as CEMIP expression status were analyzed using Cox proportional hazard models. Univariate analysis revealed significant prognostic factors including high CEMIP expression, age, anticancer treatment, and tumor stage ( $P < 0.05$ ). Multivariate analysis revealed that high CEMIP expression (hazard ratio [HR], 1.646; 95% CI, 1.224–2.214;  $P = 0.001$ ) was a significant independent prognostic factor after adjustment for potential confounding factors (Table 3). In 2 patients whose CEMIP and CA 19-9 levels were serially checked two times with tumor response evaluation, CEMIP showed a predictive function for treatment response, and it accurately predicted cancer progression compared to CA 19-9 (Supplementary Figure 1).



**Figure 4.** Correlation between CEMIP and survival in patients with pancreatic cancer using Kaplan-Meier curve. Low CEMIP levels were significantly associated with longer overall survival (median overall survival, 13.7 vs. 9.8 months,  $P = 0.0175$ ). CEMIP, cell migration-inducing hyaluronan binding protein.

| Variable           | Univariate          |         | Variable     | Multivariate        |         |
|--------------------|---------------------|---------|--------------|---------------------|---------|
|                    | HR (95% CI)         | P-value |              | HR (95% CI)         | P-value |
| CEMIP, mg/dL       |                     |         | CEMIP, mg/dL |                     |         |
| <0.429             | Ref.                |         | <0.429       | Ref.                |         |
| ≥0.429             | 1.378 (1.056–1.798) | 0.018   | ≥0.429       | 1.646 (1.224–2.214) | 0.001   |
| Age                | 1.022 (1.010–1.035) | 0.001   | Age          | 1.022 (1.009–1.036) | 0.001   |
| Sex                |                     |         |              |                     |         |
| Women              | Ref.                |         |              |                     |         |
| Men                | 1.219 (0.957–1.552) | 0.108   |              |                     |         |
| Diabetes mellitus  |                     |         |              |                     |         |
| No                 | Ref.                |         |              |                     |         |
| Yes                | 0.947 (0.729–1.229) | 0.680   |              |                     |         |
| Treatment          |                     |         |              |                     |         |
| Supportive care    | Ref.                |         | No           | Ref.                |         |
| Chemotherapy       | 0.521 (0.280–0.970) | 0.039   | Yes          | 0.243 (0.128–0.462) | <0.001  |
| Operation          | 0.134 (0.069–0.263) | <0.001  |              |                     |         |
| Location of cancer |                     |         |              |                     |         |
| Head               | Ref.                |         |              |                     |         |
| Body               | 1.372 (1.013–1.859) | 0.041   |              |                     |         |
| Tail               | 1.444 (1.049–1.986) | 0.024   |              |                     |         |
| Mixed              | 1.344 (0.883–2.045) | 0.167   |              |                     |         |
| Stage              |                     |         |              |                     |         |
| I and II           | Ref.                |         | I and II     | Ref.                |         |
| III                | 1.228 (0.862–1.748) | 0.254   | III          | 1.733 (1.137–2.639) | 0.011   |
| IV                 | 3.754 (2.807–5.020) | <0.001  | IV           | 5.310 (3.663–7.696) | <0.001  |

**Table 3.** Univariate and multivariate analysis for risk factors influencing survival in patients diagnosed with pancreatic cancer. A  $P$ -value  $< 0.02$  was considered statistically significant. We found the cut-off point to maximize the survival curve using Contal and O'Quigley's method. HR, hazard ratio; CI, confidence interval.

## Discussion

This study describes a novel biomarker, CEMIP, which may be used in combination with CA 19-9 in whole blood in the diagnosis of pancreatic cancer. CEMIP synergizes with CA 19-9 in its diagnostic function for pancreatic cancer. Combining CA 19-9 with CEMIP increased the AUROC compared with CA 19-9 alone in all groups. These results suggest the usefulness of CEMIP in combination with CA 19-9 or in situations in which the CA 19-9 level is normal.

Cancer stem cells (CSCs) are potential diagnostic and therapeutic targets because of their roles in carcinogenesis<sup>20,21</sup>. In recent research, CSC-related biomarkers have been suggested as potential reliable biomarkers for tumor detection<sup>22,23</sup>. We previously performed cDNA microarrays using adherent and sphere cells from the human pancreatic cancer cell lines Capan-1 and HPAC to evaluate the unique molecular patterns of pancreatic CSCs<sup>24</sup>. We characterized biomarkers for distinguishing pancreatic CSCs from cancer cells. cDNA microarrays with sphere and adherent cells derived from Capan-1 and HPAC cell lines revealed that the *CEMIP* gene was

upregulated in sphere cells, but not in adherent Capan-1 and HPAC cells (data not shown, Jang *et al.*, submitted). Among the candidate genes identified in this work, we were particularly interested in the *CEMIP* gene, which has not been studied extensively in CSCs<sup>21,25,26</sup>. We focused on the role of *CEMIP* in relation to CSC-related functions, especially diagnostic and prognostic function. CSC-related markers have been shown to aid early detection in cancer and have been associated with chemotherapy resistance and prognosis, because CSCs are known as tumor-initiating cells and conventional anticancer treatments do not target CSCs<sup>20–23,27–29</sup>.

Combining detection with different biomarkers has been reported to be a useful strategy in several studies, as it increases the sensitivity and specificity of each biomarker. In a combined ROC analysis using the two markers, both sensitivity and specificity reached a relatively high value in discriminating patients with pancreatic cancer from healthy individuals. Recent studies reported the effect of combined metabolic biomarkers or microRNA based molecular markers in pancreatic cancer diagnosis<sup>8,9</sup>. However, no reliable CSC-related biomarker has been established for pancreatic cancer diagnosis and prognosis combined with widely used CA 19-9. Our results suggest that the combination of CA 19-9 and *CEMIP* serum assays may offer a useful tool for the detection of pancreatic cancer.

Regarding early-stage pancreatic cancer screening, endoscopic ultrasound, magnetic resonance imaging, and computed tomography have been used in previous studies for several decades as a first choice in patients with high risk factors<sup>30–37</sup>. However, those procedures are associated with low compliance in the clinical setting due to their invasive and expensive character. Therefore, cost-effective and non-invasive biomarkers with high sensitivity and specificity are needed to enable early detection of early-stage pancreatic cancer. Thus, we evaluated the diagnostic performance of *CEMIP* in patients with stage I & II pancreatic cancer. In early-stage pancreatic cancer, combined use of *CEMIP* and CA 19-9 showed higher AUROC than CA 19-9 alone (0.95 vs. 0.85,  $P = 0.0004$ ) (Supplementary Figure 2).

There are several strengths in the present study. First, this study is the largest to investigate the role of *CEMIP* in pancreatic cancer. Second, this study is based on the finding that *CEMIP* is associated with CSCs. No previous studies have investigated the association between *CEMIP* and CSCs or the role of *CEMIP* as a CSC-related marker. Third, all clinical data and blood samples were collected from a large-scale prospective cohort of patients with pancreatic cancer.

However, there are also several limitations in this study. First, owing to restrictions in the sampling of healthy participants, there was a significant difference in age ( $P < 0.001$ ) between patients with pancreatic cancer and normal individuals (mean age, 63 vs. 49 years). Thus, the outcome might be confounded by age for healthy blood donors. However, separate analysis for cases and controls revealed no correlation between age and *CEMIP* (Supplementary Figure 3). This finding suggests that *CEMIP* might be a useful diagnostic biomarker independent of age. In addition, the number of individuals in the control groups is small compared to the number in the group with pancreatic cancer. Future prospective studies with large sample sizes including healthy individuals and patients with premalignant disease may further support our results. Lastly, the specificity of combination of *CEMIP* and CA19-9 was low (59.2%) even though the sensitivity was high (96.6%). Although the specificity of *CEMIP* was high (90.0%) to identify patients with pancreatic cancer when compared to the normal controls without benign diseases, it became low probably due to the high number of false-positive cases. Therefore, clinicians should be careful to use *CEMIP* and CA 19-9 as biomarkers for pancreatic cancer diagnosis in patients with benign disease. For these patients, other examinations including endoscopic ultrasonography and abdominal computed tomography scan might be helpful to complement the low specificity of the combination of *CEMIP* and CA 19-9.

In conclusion, serum ELISA analysis showed that *CEMIP* proteins were highly expressed in patients with pancreatic cancer compared to healthy individuals. Combined use of CA 19-9 and *CEMIP* significantly increased the sensitivity and specificity in discriminating not only patients with all stage pancreatic cancer but also patients with stage I/II pancreatic cancer from healthy individuals. Therefore, combined detection with serum CA 19-9 and *CEMIP* levels may have the potential to become a new laboratory method for the clinical diagnosis of pancreatic cancer.

## References

- Ryan, D. P., Hong, T. S. & Bardeesy, N. Pancreatic adenocarcinoma. *N Engl J Med* **371**, 2140–2141 (2014).
- Tempero, M. A. *et al.* Relationship of carbohydrate antigen 19-9 and Lewis antigens in pancreatic cancer. *Cancer Res* **47**, 5501–5503 (1987).
- van den Bosch, R. P., van Eijck, C. H., Mulder, P. G. & Jeekel, J. Serum CA19-9 determination in the management of pancreatic cancer. *Hepatology* **43**, 710–713 (1996).
- Pleskow, D. K. *et al.* Evaluation of a serologic marker, CA19-9, in the diagnosis of pancreatic cancer. *Ann Intern Med* **110**, 704–709 (1989).
- Steinberg, W. The clinical utility of the CA 19-9 tumor-associated antigen. *Am J Gastroenterol* **85**, 350–355 (1990).
- Kim, H. J. *et al.* A new strategy for the application of CA19-9 in the differentiation of pancreaticobiliary cancer: Analysis using a receiver operating characteristic curve. *American Journal of Gastroenterology* **94**, 1941–1946 (1999).
- Luo, G. *et al.* Potential Biomarkers in Lewis Negative Patients With Pancreatic Cancer. *Ann Surg* **265**, 800–805 (2017).
- Schultz, N. A. *et al.* MicroRNA biomarkers in whole blood for detection of pancreatic cancer. *JAMA* **311**, 392–404 (2014).
- Mayerle, J. *et al.* Metabolic biomarker signature to differentiate pancreatic ductal adenocarcinoma from chronic pancreatitis. *Gut*; <https://doi.org/10.1136/gutjnl-2016-312432>. (2017).
- Koga, A. *et al.* KIAA1199/*CEMIP*/*HYBID* overexpression predicts poor prognosis in pancreatic ductal adenocarcinoma. *Pancreatol* **17**, 115–122 (2017).
- Tiwari, A. *et al.* Early insights into the function of KIAA1199, a markedly overexpressed protein in human colorectal tumors. *PLoS One* **8**, e69473 (2013).
- Evensen, N. A. *et al.* Unraveling the role of KIAA1199, a novel endoplasmic reticulum protein, in cancer cell migration. *J Natl Cancer Inst* **105**, 1402–1416 (2013).
- Matsuzaki, S. *et al.* Clinicopathologic significance of KIAA1199 overexpression in human gastric cancer. *Ann Surg Oncol* **16**, 2042–2051 (2009).

14. LaPointe, L. C. *et al.* Discovery and validation of molecular biomarkers for colorectal adenomas and cancer with application to blood testing. *PLoS One* **7**, e29059 (2012).
15. Chivu Economescu, M. *et al.* Identification of potential biomarkers for early and advanced gastric adenocarcinoma detection. *Hepatogastroenterology* **57**, 1453–1464 (2010).
16. Suh, H. N. *et al.* Identification of KIAA1199 as a Biomarker for Pancreatic Intraepithelial Neoplasia. *Sci Rep* **6**, 38273 (2016).
17. Kohi, S. *et al.* KIAA1199 is induced by inflammation and enhances malignant phenotype in pancreatic cancer. *Oncotarget*; <https://doi.org/10.18632/oncotarget.15052>. (2017).
18. Edge, S. B. & Compton, C. C. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol* **17**, 1471–1474 (2010).
19. Contal, C. & O'Quigley, J. An application of changepoint methods in studying the effect of age on survival in breast cancer. *Computational statistics & data analysis* **30**, 253–270 (1999).
20. Sergeant, G., Vankelecom, H., Gremeaux, L. & Topal, B. Role of cancer stem cells in pancreatic ductal adenocarcinoma. *Nat Rev Clin Oncol* **6**, 580–586 (2009).
21. Beck, B. & Blanpain, C. Unravelling cancer stem cell potential. *Nat Rev Cancer* **13**, 727–738 (2013).
22. Klonisch, T. *et al.* Cancer stem cell markers in common cancers - therapeutic implications. *Trends Mol Med* **14**, 450–460 (2008).
23. Liu, Z. *et al.* Significance of stem cell marker Nanog gene in the diagnosis and prognosis of lung cancer. *Oncol Lett* **12**, 2507–2510 (2016).
24. Jung, D. E., Wen, J., Oh, T. & Song, S. Y. Differentially Expressed MicroRNAs in Pancreatic Cancer Stem Cells. *Pancreas* **40**, 1180–1187 (2011).
25. Chiou, S. H. *et al.* Coexpression of Oct4 and Nanog enhances malignancy in lung adenocarcinoma by inducing cancer stem cell-like properties and epithelial-mesenchymal transdifferentiation. *Cancer Res* **70**, 10433–10444 (2010).
26. Wei, F., Scholer, H. R. & Atchison, M. L. Sumoylation of Oct4 enhances its stability, DNA binding, and transactivation. *J Biol Chem* **282**, 21551–21560 (2007).
27. Hermann, P. C. *et al.* Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. *Cell Stem Cell* **1**, 313–323 (2007).
28. Li, C. *et al.* Identification of pancreatic cancer stem cells. *Cancer Res* **67**, 1030–1037 (2007).
29. Clevers, H. The cancer stem cell: premises, promises and challenges. *Nat Med* **17**, 313–319 (2011).
30. Canto, M. I. & Hruban, R. H. Diagnosis: A step closer to screening for curable pancreatic cancer? *Nat Rev Gastroenterol Hepatol* **12**, 431–432 (2015).
31. Yu, J. *et al.* Digital next-generation sequencing identifies low-abundance mutations in pancreatic juice samples collected from the duodenum of patients with pancreatic cancer and intraductal papillary mucinous neoplasms. *Gut*; <https://doi.org/10.1136/gutjnl-2015-311166>. (2016).
32. Lennon, A. M. *et al.* The early detection of pancreatic cancer: what will it take to diagnose and treat curable pancreatic neoplasia? *Cancer Res* **74**, 3381–3389 (2014).
33. Chari, S. T. *et al.* Early detection of sporadic pancreatic cancer: summative review. *Pancreas* **44**, 693–712 (2015).
34. Canto, M. I. *et al.* International Cancer of the Pancreas Screening (CAPS) Consortium summit on the management of patients with increased risk for familial pancreatic cancer. *Gut* **62**, 339–347 (2013).
35. Li, A. *et al.* MicroRNA array analysis finds elevated serum miR-1290 accurately distinguishes patients with low-stage pancreatic cancer from healthy and disease controls. *Clin Cancer Res* **19**, 3600–3610 (2013).
36. Kanda, M. *et al.* Mutant TP53 in duodenal samples of pancreatic juice from patients with pancreatic cancer or high-grade dysplasia. *Clin Gastroenterol Hepatol* **11**, 719–730 e715 (2013).
37. Shin, E. J. & Canto, M. I. Pancreatic cancer screening. *Gastroenterol Clin North Am* **41**, 143–157 (2012).

## Acknowledgements

This work was supported by the Brain Korea 21 PLUS Project for Medical Science, Yonsei University, Seoul, Korea. This research was also supported by a grant of the Korean Health Technology R&D Project (A111075), Ministry of Health & Welfare, Republic of Korea. This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI14C1324). This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (grant number: 2017R1C1B2012956).

## Author Contributions

H.S.L. and S.Y.S. designed studies; H.S.L., C.Y.J., S.A.K., S.B.P. and D.E.J. performed experiments and prepared all the figures; H.S.L., S.A.K., and S.Y.S. analyzed data; B.O.K., and H.Y.K. performed statistical analysis; H.S.L., M.J.C., J.Y.P., S.B., S.W.P., and S.Y.S. wrote the paper. All authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-21823-x>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018