



Perspective

Protein Biomarkers for Early Detection of Pancreatic Ductal Adenocarcinoma: Progress and Challenges

Alex Root ¹, Peter Allen ^{2,3}, Paul Tempst ^{1,*}^(D) and Kenneth Yu ^{3,4,*}

- ¹ Molecular Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA; roota@mskcc.org
- ² Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA; allenp@mskcc.org
- ³ Department of Medicine, Weill Cornell Medical College, New York, NY 10065, USA
- ⁴ Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
- * Correspondence: p-tempst@ski.mskcc.org (P.T.); yuk1@mskcc.org (K.Y.); Tel.: +1-212-639-8923 (P.T.); +1-646-888-4188 (K.Y.)

Received: 29 November 2017; Accepted: 3 March 2018; Published: 7 March 2018

Abstract: Approximately 75% of patients with pancreatic ductal adenocarcinoma are diagnosed with advanced cancer, which cannot be safely resected. The most commonly used biomarker CA19-9 has inadequate sensitivity and specificity for early detection, which we define as Stage I/II cancers. Therefore, progress in next-generation biomarkers is greatly needed. Recent reports have validated a number of biomarkers, including combination assays of proteins and DNA mutations; however, the history of translating promising biomarkers to clinical utility suggests that several major hurdles require careful consideration by the medical community. The first set of challenges involves nominating and verifying biomarkers. Candidate biomarkers need to discriminate disease from benign controls with high sensitivity and specificity for an intended use, which we describe as a two-tiered strategy of identifying and screening high-risk patients. Community-wide efforts to share samples, data, and analysis methods have been beneficial and progress meeting this challenge has been achieved. The second set of challenges is assay optimization and validating biomarkers. After initial candidate validation, assays need to be refined into accurate, cost-effective, highly reproducible, and multiplexed targeted panels and then validated in large cohorts. To move the most promising candidates forward, ideally, biomarker panels, head-to-head comparisons, meta-analysis, and assessment in independent data sets might mitigate risk of failure. Much more investment is needed to overcome these challenges. The third challenge is achieving clinical translation. To moonshot an early detection test to the clinic requires a large clinical trial and organizational, regulatory, and entrepreneurial know-how. Additional factors, such as imaging technologies, will likely need to improve concomitant with molecular biomarker development. The magnitude of the clinical translational challenge is uncertain, but interdisciplinary cooperation within the PDAC community is poised to confront it.

Keywords: pancreatic ductal adenocarcinoma; early detection; biomarkers; blood test; ELISA; mass spectrometry; circulating DNA; thrombospondin; *CA19-9*; KRAS

1. Unmet Needs for Early Detection of Pancreatic Ductal Adenocarcinoma

Pancreatic ductal adenocarcinoma (PDAC) is an aggressive, chemorefractory and recalcitrant cancer and the number of deaths it causes has recently surpassed breast cancer [1]. PDAC is often detected when it is in a relatively advanced stage with only approximately 1 in 4 patients having resectable tumors [2]. Early detection when PDAC is at Stage I or II can provide a window of opportunity when the disease can be stopped [3]. The experience in Japan with pancreatic cancer registry shows that early detection of PDAC improves overall survival [4]. Therefore, development of an early diagnostic test is a major unmet need. A promising source for an early diagnostic test is blood-based molecular biomarkers, which include proteins, nucleic acids, autoantibodies, aberrantly glycosylated antigens, exosomes, circulating tumor cells, and metabolites or panels thereof. These ideal, non-invasive biomarkers would also be universally present in advanced pre-cancerous lesions (PanIN2-3, IPMN with high-grade dysplasia, carcinoma in situ), specific, sensitive, inexpensive, rapid, and practical to perform [5]. This article aims to describe recent progress and remaining challenges for this research-in-translation and for protein biomarkers more generally.

2. Development of Molecular Biomarkers

Translating promising research on biomarkers to clinically useful tests is a complex interdisciplinary problem with few successes [6]. In an insightful perspective article Rifai and colleagues describe the pipeline for development of molecular biomarkers consisting of six major steps: discovery, qualification, verification, research assay optimization, clinical validation, and regulatory approval & commercialization [7]. In our opinion, it remains the best roadmap to guide this multiyear process and we have illustrated it in Figure 1, including challenges and pitfalls. In discovery (step 1), the objective is to profile model systems or patient samples to nominate candidates with differential abundance between PDAC and normal, PDAC and chronic pancreatitis, or PDAC and other non-adenocarcinoma neoplasms of the pancreas. The vast majority of published biomarker research ends at this stage and, given the thousands of studies already published, is arguably useless unless it proceeds to subsequent stages [7]. Qualification (step 2) tests the detectability and discriminatory power of candidate biomarkers in human samples, usually bodily fluids, e.g., serum or plasma, and compares cases to controls using targeted technologies such as PCR or LC-MS/MS with multiple reaction monitoring (MRM) or parallel reaction monitoring (PRM). Verification (step 3) involves assessing the sensitivity and specificity of the candidates in large numbers of samples, typically 100 or more. Research assay optimization (step 4) for nucleic acid, protein, and metabolite candidate biomarkers requires development and characterization for immuno- or mass spectrometry-based assays. Research assay optimization will likely include a second-step where individual assays are optimized together as a panel [8]. Clinical validation (step 5) requires design of a prospective trial with regulatory approval in mind. Pepe and colleagues provided a rigorous system to guide study design with their PRoBE guidelines [9]. Regulatory approval may occur in concert with commercialization (step 6).

The search for biomarkers is motivated by the knowledge that early detection of PDAC will be clinically useful because timely surgical resection can lead to a cure [4]. Cohen and colleagues note that a clinically useful screening test is likely to uncover a greater proportion of earlier-stage tumors because the individuals are asymptomatic and have smaller tumors, which are easier to cure than the typical patients presenting with PDAC today [10]. Additionally, from a theoretical evolutionary dynamics perspective, the tumor population size is a critical determinant of the probability it will evolve resistance in response to therapy and, therefore, early detection of consequently smaller tumors has tremendous implications for eradication by neo-adjuvant therapy and surgery [11].

There is vigorous debate in the community about how feasible early detection of PDAC is for different risk groups. Pannala and colleagues argued "since the incidence of pancreatic cancer is low, screening for asymptomatic cancer in the general population will not be feasible", and instead argue for a two-tiered screening strategy of stratifying patients into high-risk cohorts based on family history, recent-onset diabetes, or other factors to be determined [12]. We note that in some cases screening for rare diseases is feasible and clinically useful. Phenylketonuria has low incidence, but screening for it is a major medical accomplishment [13], so rarity alone is not dispositive. Panalla and colleagues describe concrete numbers:

"The age-adjusted incidence of pancreatic cancer in subjects \geq 50 years of age is 38/100,000. If a test with 99% sensitivity and 99% specificity for pancreatic cancer is used to screen 100,000 subjects \geq 50 years of age, the test would identify nearly all pancreatic cancers in



the population screened (n = 37). However, the test would also falsely identify nearly a 1000 subjects as having pancreatic cancer." [12]

Figure 1. Illustration of the biomarker pipeline. Discovery of biomarkers typically utilizes mass spectrometry, nuclear acid sequence analysis, or ELISAs. In Qualification and Verification the objective is to find biomarkers or a panel of biomarkers that can separate cancer from benign disease and normal controls. During Assay Optimization the linear range of the assay is rigorously assessed and improved, if possible. There must be a representative diversity of patient samples chosen for Assay Optimization, so that the full spectrum of sample heterogeneity that may be ultimately encountered in the clinic is properly optimized for. For Validation, a prospective clinical trial is conducted and cost effectiveness is further assessed. Commercialization may proceed in concert with Validation and requires a demonstration of clinical utility. As the pipeline progresses, there is a sharp decrease in the number of biomarker candidates analyzed and a corresponding increase in the number of patient samples analyzed.

Pannala and colleagues are correct to point out that even a diagnostic test with excellent specificity and sensitivity seems to have an impossible, high bar. However, we would add that evaluation of the utility of a biomarker is complex and requires cost-effectiveness analysis that includes a holistic evaluation, including what happens subsequently to those individuals who take the test. Of particular importance for rare diseases is the cost of false-positives. If the test itself is inexpensive, and the consequence of a false positive is a non-invasive imaging test, it is a less costly outcome then if the patient receives an unnecessary invasive biopsy. Therefore, it is an oversimplification to merely require that a biomarker achieves cutoffs for specificity and sensitivity. Careful cost-effectiveness analysis of the test itself and its consequences and alternatives is needed to assess clinical utility. The frameworks of decision analysis and economic evaluation were developed to better evaluate clinical utility than accuracy measures, such as sensitivity and specificity [14]. Andrew Vickers provided a refinement to simplify decision analysis with decision curve analysis [15], noting that traditional accuracy measures such as sensitivity, specificity and positive predictive value are useful in the early stages of biomarker development [15]. Ghatnekar and colleagues provided a detailed examination of these advanced cost-effectiveness concepts, using quality-adjusted life years (QALY), incremental cost per QALY gained (ICER), and cost-effectiveness acceptability curve (CEAC) [16]. They modeled these concepts applied to PDAC biomarker panels, concluding that for high-risk patients a serum biomarker signature was highly desirable [16].

The most probable path for achieving clinical usefulness in the near-term is to follow the two-tiered screening strategy of Pannala and colleagues, whereby firstly, high-risk patients are identified and then screened [12]. High-risk patients comprise 10–15% of PDAC and include those with pancreatitis, intraductal papillary mucinous neoplasms, recent onset diabetes, *CDKN2A*, *BRCA1/2*, *PALB2*, *STK11*, *ATM*, *CTFR* germline mutations, and family history with two or more first-degree relatives having PDAC [17,18]. Given that screening high-risk patients is the intended use, it is an important question as to whether these patients behave as a distinct subset with different biomarker panel performance than non-high-risk cases. Norris and colleagues compared the genomic progression of familial vs. sporadic PDAC and concluded that it is the same [19]. Therefore, both sporadic and familial cases can be used to develop biomarkers, even though the intended use is screening high-risk cases. This is significant because high-risk cases are rarer and limiting patient accrual to this subset would make it much more difficult to reach the hundreds of cases needed for statistical power. The test itself is likely to be a panel of biomarkers [8]. An important question for the field is: how high should numerical cutoffs for area-under-curve, sensitivity, and specificity be for high-risk cohorts?

We recommend that efforts to find biomarker panels for non-high-risk patients continue even though they do not represent the population for intended use because: (1) the genomic progression of PDAC appears to be the same whether patients have familial or sporadic PDAC; therefore, studying both groups together is pragmatic for patient recruitment; (2) the definition of "high-risk" continues to change as more is learned about gene-disease and other risk factors; (3) candidate biomarker levels could be useful to stratify patients into high risk groups for surveillance, and therefore, the search for early detection biomarkers in non-high-risk cases could help predict risk; (4) improvements in sequencing and mass spectrometry technologies may uncover truly PDAC-specific signatures with >99% sensitivity and specificity; such revolutionary developments have happened before, as occurred with newborn screening for in-born errors of metabolism; (5) the enormous potential cost savings of an early detection test for PDAC more than compensates for the apparently low probability that a general population screening test can be developed—in other words, the risk is worth the reward.

3. Recent Advances in Protein Biomarkers for Early Detection of PDAC

The research community has devoted much time and resources to developing biomarkers for early detection of PDAC with better performance than *CA19-9* [3]. Current clinical practice uses an ELISA for *CA19-9*, which is a carbohydrate found on multiple carrier proteins, whose identities may themselves have diagnostic significance [20]. However, *CA19-9* is not detectable in 5–10% of patients who lack fucosyltransferase activity due to germline variants [21]. *CA19-9* is useful for monitoring response to therapy, but not very useful as an early detection biomarker [20]. Whether it should be part of a biomarker is uncertain. However, it is routinely measured as part of clinical care of research cohorts, and therefore, should be included in biomarker studies.

Biomarkers for diagnosis, treatment prediction, and prognosis of PDAC were reviewed by Chari and colleagues [5], Bunger and colleagues [22], Loosen and colleagues [23], Gallego and colleagues [24], and exosome biomarkers by Li and colleagues [25]. Four key studies nominating primarily protein diagnostic biomarkers of particular—albeit highly subjective—interest to us, which report outperforming *CA19-9* alone, are briefly summarized in Box 1, and their performance characteristics are shown in Table 1. We note at the outset that most are multivariate panels including *CA19-9*.

Box 1. Recently reported proteomic biomarker panels with better performance than CA19-9.

Cohen 2017 [10]: Found that combining ctDNA testing for KRAS mutations in combination with 4 plasma proteins (*CA19-9, CEA, HGF, OPN*) outperformed *CA19-9* alone in discriminating PDAC from normal controls, chronic pancreatitis, and other benign pancreatic diseases. A disappointing 64% sensitivity begs the question as to whether including additional biomarkers or optimizing the assays may improve performance. Evaluation in an independent test set is desirable.

Capello 2017 [21]: Began with 17 protein biomarker candidates from previous studies and validated several that can distinguish PDAC from controls and chronic pancreatitis (*TIMP1*, *LRG1*, *REG3A*, *IGFBP2*, *COL18A1*, *TNFRSF1A*), finding *TIMP1* + *LRG1* + *CA19-9* outperformed *CA19-9* alone. Evaluation in an independent test set is desirable.

Kaur 2017 [26]: Performed an in-depth analysis of the literature and observed that *MUC5AC* has favorable biomarker properties: secretion, high over-expression in PanIN, and numerous epitopes. They developed an in-house ELISA against *MUC5AC*, finding that *MUC5AC* + *CA19-9* outperforms *CA19-9* alone in two, large independent validation cohorts.

Kim 2017 [27]: Began with a state-of-the-art cell reprogramming model of PDAC and sophisticated systems biology network analysis to nominate thrombo-spondin-2 (*THBS2*); when combined with *CA19-9* it outperformed *CA19-9* in discriminating PDAC from controls and chronic pancreatitis. The two biomarker panel was evaluated in a discovery phase and two validation phases, however, the number of Stage I/II patients was 0 in the discovery phase, and 7 Stage I and 34 Stage II in validation set A.

Table 1. Summarized performance of biomarker candidates. Listed here are published performance metrics for cancer vs. healthy or benign controls and cancer vs. chronic pancreatitis in independent validation datasets when available, across all stages reported. Studies do not always make these basic statistics available in an easy-to-read table and not all numbers are directly comparable.

Study	n (Training; val1; val2) Training Stage I/II Val. A Stage I/II Val. B Stage I/II Val. C Stage I/II	Biomarkers(s)	Cancer vs. Control AUC, SN, SP	Cancer vs. CP AUC, SN, SP
Poruk 2013 [28]	220; 0; 0 ^ 7/35 Stage I/II na na	CA19-9 OPN TIMP1 CA19-9 + OPN + TIMP1	92, 84, 88 *.^ 72, na, na *.^ 77, na, na *.^ na, 87, 91 *.^	*
Radon [29]∇	134; 58 4/14 Stage I/II Unclear	LYVE1 REG1A TFF1 panel + creatine + age panel + plasma CA19.9	84, 68, 92; na, na, na 75, 75, 69; na, na, na 70, 79, 53; na, na, na 90, 82, 89; 93, 80, 77 na, na, na; 97, 88, 96	73, 77, 62; na, na, na 72, 75, 70; na, na, na 65, 77, 56; na, na, na 83, 86, 67; 85, 100, 50 na, na, na; 87, 75, 94
Kaur 2017 [26]	346; 94; 321 70 Stage I/II Unclear Unclear	МИС5АС СА19-9 МИС5АС + СА19-9	84,70,83; 70,68,73; 74,65,83 [§] 57,48,67 [§] ; na, na, na; na, na, na 86,72,85 [§] ; na, na, na; na, na, na, na	*
Capello 2017 [21]	121; 30; 142; 35~ Unclear 10 Stage I/II 42 Stage I/II 21 Resectable	CA19-9 TIMP1 LRG1 TIMP1 + LRG1 + CA19-9 ("OR" rule)	88, 73, 23~ 81, 41, 50~ 85, 43, 25~ 96, 85, 67~	82, 29, 24~ 73, 22, 33~ 68, 11, 12~ 89, 45, 54~
Kim 2017 [27]	20; 189; 537 0 Stage I/II 7/34 Stage I/II 4/37 Stage I/II	CA19-9 THBS2 CA19-9 + THBS2	85, 69, 100; 58, 78, 99 84, 33,96; 58, 94, 75 96, 74, 96; 76, 88, 93	77, na, na; 82, na, na 73, na, na; 73, na, na 84, na, na; 87, na, na
Cohen 2017 [10]	na; 403; 0 na 29/102 Stage I/II	ctDNA CA19-9 CEA + HGF + OPN ctDNA + CA19-9 ctDNA + CEA + HGF + OPN CA19-9 + CEA + HGF + OPN ctDNA+CA19-9+CE+ HGF + OPN	na, 30, na na, 49, na na, 18, na na, 60, na na, 42, na na, 54, na na, 64, 99.5	na, na, na na, na, na na, na, na na, na, na na, na, na na, na, na na, na, na

n total cases and controls, *val1/2* validation datasets, *SN* sensitivity, *SP* specificity, *CP* chronic pancreatitis, *na* not available). * control and chronic pancreatitis combined, ^ performance on training set, ~3rd independent validation dataset used and performance reported for combined validation datasets. [§] Early pancreatic cancer vs. healthy controls. ∇ Stage I/II vs. healthy control or chronic pancreatitis.

Molecular profiling of PDAC model systems is a widely used strategy for biomarker discovery. Cell lines, mouse models, and organoids can satisfactorily recapitulate disease progression and be more amenable to intervention and analyses in the laboratory than precious samples from human patients. In 2008 Faca and colleagues took an innovative approach by using *Pdx1-Cre Ink4a/Arf^{dox/lox}* and *Kras^{G12D} Ink4a/Arf^{dox/lox}* genetically engineered mice in order to sample blood from normal, early PanIN, advanced PanINs, locally advanced cancer, and metastatic cancer [30]. Using a tour-de-force proteomics approach to identify biomarkers directly in plasma they followed a major protein immune-depletion strategy, isotopic labeling, and extensive pre-fractionation before identifying 1442 unique proteins with LC-MS/MS. Additional candidates were included with mRNA analysis; absence in the mouse liver proteome; high expression ratio PDAC/PanIN; not annotated for acute-phase reactants or coagulation; and having a human orthologue—resulting in 45 proteins—a handful of which had commercially available antibodies or ELISA kits. These were quantified in mouse pancreas tissue or blood and researched for previous analysis in published studies. Subsequently 10 proteins were tested in 30 human PDAC sera, 20 matched controls, and 10 chronic pancreatitis cases. Performance of the markers distinguishing PDAC from controls and chronic pancreatitis revealed statistically significant biomarker candidates *ALCAM1*, *TIMP1*, *ICAM1*, *REG3*, *IGFBP4* panel plus *CA19-9*. The authors also note a striking concordance between mouse and human markers, with 8 out of 9 concurring.

Given the plethora of diagnostic biomarker discovery studies already published it can be productive and tremendously cost-effective to analyze the literature and combine reported candidate markers in a panel, instead of performing a costly discovery step 1, and thereby begin at step 2 (qualification). Poruk and colleagues employed this strategy and assessed the utility of secreted proteins *OPN* (osteopontin) and *TIMP1* (tissue inhibitor of metalloproteinase 1) plus *CA19-9* as diagnostic or prognostic biomarkers using commercially available ELISA kits [28]. Serum from 220 subjects was analyzed and an iterative classification tree analysis determined cut-points of *CA19-9* above 37U/mL followed by *OPN* 25 ng/mL and *CA19-9* below 37U/mL further split by *TIMP-1* level of 411 ng/mL. Panel performance plus *CA19-9* is summarized in Table 1. Testing performance of these markers in additional validation datasets is highly desirable.

Comprehensive genomic analyses of PDAC tissue and cell lines is leading to a new understanding of disease biology [31] and overexpressed genes are a sensible source for biomarker candidates. Kaur and colleagues began with this insight and analyzed over-expression data for the most highly expressed mucin gene in PDAC relative to benign controls, nominating *MUC5AC*. They also observed that it has additional desirable biomarker attributes including high expression in early disease stage (PanIN), secreted protein, and containing multiple epitopes [26]. They developed a highly sensitive sandwich ELISA (University of Nebraska Medical Center) using a monoclonal antibody provided by Jacques Bara (INSERM). Diagnostic training set samples were obtained from University of Pennsylvania Medical Center; validation set 1 was obtained from The Mayo Clinic; validation set 2 from University of Pennsylvania Medical Center; and tissue microarrays from Anirban Maitra (Johns Hopkins). This study is to be commended for truly community-wide collaborative efforts. Diagnostic performance of *MUC5AC*, *CA19-9* and combination panel is shown in Table 1.

In a complementary approach to developing protein biomarkers in blood (serum or plasma), urine has also been used as source for biomarkers. Radon and colleagues analyzed 18 PDAC, chronic pancreatitis, and normal controls by mass spectrometry [29]. Data analysis revealed LYVE1, REG1A/B, and TFF1 having discriminatory power in both genders. Presence of full-length proteins was confirmed by western blot in urine. ELISA assays for 488 urine samples were done with performance characteristics shown in Table 1. Data was randomly partitioned into training and test sets (70%/30%). Cutpoint optimization was performed on the training set using the R pROC software package, adjusted by creatine levels and age into a 5-part panel logistic regression model. Testing in PDAC tissues using these urine biomarkers in combination with blood-based biomarkers or miRNA urine biomarkers [32] is highly desirable.

In-depth analysis of the literature and datasets is required to put new information in context; performing this scholarly work, termed "context placement", can help make clinical research more

useful [33]. Capello and colleagues revisited the data in Faca 2008 [30] and performed an analysis of the literature to arrive at 17 candidates plus *CA19-9* using commercially available ELISA kits for 16 proteins and developing their own ELISA for 1 protein. Seven candidates passed qualification (step 2) and verification (step 3) in multiple, independent datasets that considerably advanced towards completion of research assay optimization (step 4) and clinical validation (step 5) [21]. Performance of the biomarker candidates is shown in Table 1. A particularly noteworthy statistical reporting advance was made in this study, and involves finding and reporting: sensitivity at 95% specificity and specificity at 95% sensitivity, where most studies only report maximum sensitivity and specificity; and evaluation of biomarkers and panels in 3 totally independent validation datasets where most studies only report a single validation set.

The development of organoids recapitulating PDAC is a major recent achievement and one of their numerous good uses is biomarker development [2]. Of comparable magnitude, the network biology paradigm has advanced analysis of biological systems to consider how biomolecules act multilaterally to achieve complex phenotypes rather than a reductivist view of single molecules acting unilaterally [34]. Putting both of these advances together, Kim and colleagues used the PanIN organoid model "10-22 cell-derived" and cross-referenced their 107 secreted proteins with: low abundance in the healthy human plasma proteome; RNA-seq datasets; commercial availability of ELISAs; and presence in the integrated networks for TFG-beta and integrin signaling, which drive PDAC [27]. This resulted in an investigation of MMP2, MMP10 and THBS2 (thrombospondin 2) plus CA19-9. Initial assessment revealed the discriminatory power of THBS2, and this was subsequently validated in two phases with performance characteristics of the two-panel marker shown in Table 1. This study is also notable for its statistical designs: adherence to PRoBE guidelines; using 3 different lot numbers for ELISA kits; blinding experimenters with de-identified samples; performing ELISAs in multiple laboratories; and comparison of PDAC vs. multiple benign diseases. While a large number of samples were tested in validation sets A and B, most samples were Stage III/IV, so further testing of this panel in Stage I/II cohorts is desirable in order to establish performance in the intended use of early detection of Stage I/II PDAC.

Mutations in *KRAS* are highly recurrent in PDAC and drive tumorigenesis from the earliest stages, making them a highly desirable target biomarker for early detection [35]. Cohen and colleagues combined measurement of *KRAS* mutations in ctDNA with several protein biomarkers *CEA*, *HGF*, *midkine*, *OPN*, and *prolactin* plus *CA19-9*. Their best performing combination assay consisted of *ctDNA*, *CA19-9*, *CEA*, *HGF*, and *OPN* with performance characteristics summarized in Table 1. They also found that the panel had prognostic value beyond clinical and histopathology in current use—patients with poorer survival were more likely to have a positive test. Overall, this study attempted to demonstrate that combining nucleic acid and protein candidate biomarkers is a promising approach; however, further research is needed to establish this.

There are a number of take-aways from the studies highlighted in Box 1. Commonalities include sophisticated use of model systems or existing literature for candidate discovery that advance the field beyond the use of traditional cell lines or the difficulties of direct discovery in plasma. Most studies used plasma or serum rather than bodily fluids such as urine, pancreatic juice, bile, or ascites. The validation method of choice is almost always ELISAs and frequently the commercial availability of ELISA kits is a make-it-or-break-it filter for advancing a candidate through the biomarker pipeline. Diversity in approaches among studies included their measurement technologies, the number and size of validation datasets, the specific reporting of statistics, and the case-control comparisons. All studies found improved performance with a multivariate panel including *CA19-9* over any one marker alone. Most studies assume that *CA19-9* is "an anchor marker", and therefore will increase performance with candidate markers and in combination [21,36]; however, this assumption should be questioned, and we recommend authors compute performance for all combinations of all biomarker candidates. Despite the promising performance of these studies shown in Table 1, significant challenges remain.

4. Biomarker Challenges

4.1. Challenges in Discovery, Qualification, and Verification

There have been numerous papers published with PDAC candidate biomarkers, even though this research is difficult and time-consuming [23,24], because the human plasma proteome contains proteins spanning seven orders of magnitude in concentration, the heterogeneity of normal individuals, and disease heterogeneity that all present enormous analytical challenges [37]. Arguably, the field is also overwhelmed with published candidates. Harsha and colleagues examined more than 5000 published articles on PDAC and performed meta-analysis to filter towards a list of 60 proteins that antibodies were developed against and raised the question of whether any more biomarker discovery studies should be performed in the field [38]. In our opinion biomarker discovery studies should continue provided they proceed through the first 3 steps of the pipeline (discovery, qualification, verification) with adequate numbers of samples, because technological and methodological advances have increased assay sensitivity such that low abundant biomarkers that were invisible in past discovery scans may now be discovered. Similarly, new model systems including organoids and genetically engineered mice should continue to be explored.

An issue for discovery studies is that the methods and reporting of statistics is not uniform, which makes comparison difficult and necessitating meta-analyses and head-to-head comparisons. Studies should also report performance by stage (IA, IB, II, III, IV), because the intended use is detection of stages I/II. Guidelines for studies have been proposed [9]; however, implementing them in practice is challenging [39]. Perhaps less widely known but equally significant are the statistical constraints on the development of multivariate classifiers from large numbers of variables and relatively small sample sizes that lead to tradeoffs in the stability of the variables chosen and classification error, termed "the curse of dimensionality" [40]. Best practices in biomarker statistics require researchers to avoid over-fitting, cherry-picking, and other analysis pitfalls [9]. There has been enormous progress in methodology, and biomarker panels appear to be the best way forward [8].

4.2. Challenges in Assay Optimization, Validation, and Commercialization

Given that there are now multiple reported early detection panels for PDAC with area-under-curve, sensitivity and specificity above 0.75, and commercially available ELISAs, what challenges must be overcome to advance along the biomarker pipeline? We believe top priorities for the field are: (1) optimize assays and assay panels for the top candidates; (2) assess the top candidates as panels in independent cohort sets; (3) propose threshold values for accuracy metrics (AUC, sensitivity, specificity, PPV, NPV) for the intended-use of screening high-risk groups. What can we learn from past experience with biomarker failures?

In a blistering critique of the biomarker field, Scott Kern argues for a wholesale change in practices and a recognition of why biomarkers fail including "lack of statistical significance, hidden structure in the source data, a technically inadequate assay, inappropriate statistical methods, unmanageable domination of the data by normal variation, implausibility, deficiencies in the studied population or investigator system, and its disproof or abandonment for cause by others" [41]. Ioannidis categorized biomarker failures into 4 categories: clinical reversal (type A), validation failure (type B), non-optimized clinical translation (type C), and promotion despite non-promising evidence (type D) [42]. It is important to decide whether top candidates are being abandoned for cause and if so, state this clearly in a publication. Is there a validation failure (type B)? Otherwise, without an attempt at completing the biomarker pipeline the markers will fail due to non-optimized clinical translation (type C failure.).

Past experience has taught us that the vast majority of biomarker research does not advance past step 3 verification, despite reportedly excellent performance of biomarker candidates [6,7]. How do we improve this bottleneck? A priority for funding should be to incentivize assay optimization and facilitate sample repositories for scaling up. We recommend that leadership in the United States from the Early Detection Research Network (NCI) continue to help facilitate this, in particular by providing

funding for assay optimization and validation. Even though commercially available assays exist, their performance characteristics may be inadequate for approval under FDA requirements. Assays may have high batch-to-batch variation, poor signal-to-noise, inadequate limit-of-detection, or a narrow linear range.

Testing panels of top candidates in independent validation datasets is the best way see whether they really work [41]. Difficulties in obtaining patient samples for analysis have been ameliorated by the creation of the Mayo Clinic Prospective Resource for Biomarker Validation and Early Detection of Pancreatic Cancer, which promises only a page of paperwork [43]. Another source for patient cohorts is surveillance studies of high-risk populations, where surveillance is typically imaging-based [18,44,45]. Consortia such as the National Familial Pancreatic Tumor Registry also seek to identify and study these patients. Large surveillance studies have been conducted on high-risk cohorts with demonstrated survival benefits [18]. Potjer and colleagues examined performance of a serum-based proteomic panel for early detection in a high-risk cohort undergoing imaging-based surveillance for PDAC [46]. This study design is one way forward because it incorporates intended use of screening high-risk patients, as well as imaging. The difficulty is that high-risk cohorts are smaller than all PDAC cases. As was stated earlier, sporadic and familial pancreatic cancers share common genomic driver events [19]. Scaling up may require creation of another prospective resource for biomarker validation in addition to Mayo Clinic's.

There are also socio-economic challenges to translate biomarkers. For example, given that there are already commercially developed ELISA kits for a number of promising candidate markers and these markers typically perform best as part of a panel, it is natural to ask whether the regulatory environment is adequately incentivizing development of a marker panel consisting of diagnostic assays from multiple companies. The government may very well need to incentivize collaboration among companies to optimize and validate assay panels. Experience in promoting collaboration among academia, foundations, and industry in developing novel therapies was recently reviewed by Ramsey and colleagues and contains useful lessons that may be applied to translating biomarker research [47]. Stern and colleagues discuss incentives for biomarker development, recommending that NIH target and increase funding [48]. They also mention the FDA's Accelerated Approval Program [48]. To better incentivize biomarker development, collaboration among industry, and recognize the high risks involved in this endeavor, we boldly recommend an additional five years of patent protection for early detection tests.

5. Outlook

Progress has been achieved discovering and assessing biomarkers for early detection of PDAC. The field has nominated several panels with sensitivity and specificity above 0.75, with most panels not tested in independent validation sets, and most combinations of candidate markers not yet assessed. In our opinion, investment in assay optimization and independent validation studies should be a moonshot priority. There are significant challenges remaining including: optimizing assay and panel performance; scaling up validation studies to include hundreds of cases and controls; variability and hidden data structure that result in poor generalizability of panel performance in independent cohorts; and furthering collaboration among interdisciplinary teams from research, industry, and government. It is also important for the field to learn from previous experience, so as not to repeat mistakes [41]. Surveillance studies of high-risk individuals that use imaging can incorporate blood sampling for biomarker development, but this adds complexity and the necessity for teams with multiple areas of expertise. Case-control validation cohorts incorporating all PDAC risk groups with an intended use of a two-tiered strategy of identifying and screening high-risk patients is the most probable path forward. The initial application of biomarkers is likely to be in the high-risk population. It has nothing to do with biology, but everything to do with statistics. For a given sensitivity and specificity, the performance of a biomarker is going to vary greatly based on the disease incidence in the population screened. Once a biomarker panel is developed, it will first be used to monitor high-risk patients, followed by

confirmatory tests (imaging, endoscopy, etc.). If the test is really good, then it could be used more broadly in populations where the incidence is lower (for example, older patients without other risk factors, etc.). Despite many challenges, the progress made to-date in finding biomarkers for early detection of PDAC provides optimism and invigoration to the field.

Acknowledgments: We thank Arpi Nazarian and Kevin Lawlor for help with the figure and critical reading of the manuscript. The authors are supported by National Institutes of Health grants R21 CA174537 (Paul Tempst); R01 CA208401 (Paul Tempst; Kenneth H. Yu); R01 CA194321 and R01 CA202762 (Kenneth H. Yu); R01 CA182076 and U01 CA152653 (PA); Cancer Center Support Grant P30 CA008748; and a Thompson Foundation Grant (Paul Tempst; Kenneth H. Yu). We thank the reviewers for helpful comments and critiques, which resulted in substantial intellectual contributions.

Author Contributions: Alex Root, Paul Tempst and Kenneth H. Yu conceived, designed and wrote the paper. Alex Root created the figure. All authors critically revised the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. SEER. Cancer Stat Facts: National Cancer Institute. 2017. Available online: https://seer.cancer.gov/statfacts/ (accessed on 29 November 2017).
- 2. Kleeff, J.; Korc, M.; Apte, M.; La Vecchia, C.; Johnson, C.D.; Biankin, A.V.; Neale, R.E.; Tempero, M.; Tuveson, D.A.; Hruban, R.H.; et al. Pancreatic cancer. *Nat. Rev. Dis. Prim.* **2016**, *2*, 16022. [CrossRef] [PubMed]
- 3. Lennon, A.M.; Wolfgang, C.L.; Canto, M.I.; Klein, A.P.; Herman, J.M.; Goggins, M.; Fishman, E.K.; Kamel, I.; Weiss, M.J.; Diaz, L.A.; et al. The early detection of pancreatic cancer: What will it take to diagnose and treat curable pancreatic neoplasia? *Cancer Res.* **2014**, *74*, 3381–33819. [CrossRef] [PubMed]
- Matsuno, S.; Egawa, S.; Fukuyama, S.; Motoi, F.; Sunamura, M.; Isaji, S.; Imaizumi, T.; Okada, S.; Kato, H.; Suda, K.; et al. Pancreatic Cancer Registry in Japan: 20 years of experience. *Pancreas* 2004, 28, 219–230. [CrossRef] [PubMed]
- 5. Chari, S.T.; Kelly, K.; Hollingsworth, M.A.; Thayer, S.P.; Ahlquist, D.A.; Andersen, D.K.; Batra, S.K.; Brentnall, T.A.; Canto, M.; Cleeter, D.F.; et al. Early detection of sporadic pancreatic cancer: Summative review. *Pancreas* **2015**, *44*, 693–712. [CrossRef] [PubMed]
- Paulovich, A.G.; Whiteaker, J.R.; Hoofnagle, A.N.; Wang, P. The interface between biomarker discovery and clinical validation: The tar pit of the protein biomarker pipeline. *Proteom. Clin. Appl.* 2008, *2*, 1386–1402. [CrossRef] [PubMed]
- 7. Rifai, N.; Gillette, M.A.; Carr, S.A. Protein biomarker discovery and validation: The long and uncertain path to clinical utility. *Nat. Biotechnol.* **2006**, *24*, 971–983. [CrossRef] [PubMed]
- 8. Borrebaeck, C.A. Precision diagnostics: Moving towards protein biomarker signatures of clinical utility in cancer. *Nat. Rev. Cancer* 2017, *17*, 199–204. [CrossRef] [PubMed]
- Pepe, M.S.; Feng, Z.; Janes, H.; Bossuyt, P.M.; Potter, J.D. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: Standards for study design. *J. Natl. Cancer Inst.* 2008, 100, 1432–1438. [CrossRef] [PubMed]
- Cohen, J.D.; Javed, A.A.; Thoburn, C.; Wong, F.; Tie, J.; Gibbs, P.; Schmidt, C.M.; Yip-Schneider, M.T.; Allen, P.J.; Schattner, M.; et al. Combined circulating tumor DNA and protein biomarker-based liquid biopsy for the earlier detection of pancreatic cancers. *Proc. Natl. Acad. Sci. USA* 2017, 114, 10202–10207. [CrossRef] [PubMed]
- Bozic, I.; Reiter, J.G.; Allen, B.; Antal, T.; Chatterjee, K.; Shah, P.; Moon, Y.S.; Yaqubie, A.; Kelly, N.; Le, D.T.; et al. Evolutionary dynamics of cancer in response to targeted combination therapy. *eLife* 2013, 2, e00747. [CrossRef] [PubMed]
- 12. Pannala, R.; Basu, A.; Petersen, G.M.; Chari, S.T. New-onset diabetes: A potential clue to the early diagnosis of pancreatic cancer. *Lancet Oncol.* **2009**, *10*, 88–95. [CrossRef]
- 13. Groselj, U.; Tansek, M.Z.; Battelino, T. Fifty years of phenylketonuria newborn screening—A great success for many, but what about the rest? *Mol. Genet. Metab.* **2014**, *113*, 8–10. [CrossRef] [PubMed]
- 14. Prosser, L.A.; Grosse, S.D.; Kemper, A.R.; Tarini, B.A.; Perrin, J.M. Decision analysis, economic evaluation, and newborn screening: Challenges and opportunities. *Genet. Med.* **2012**. [CrossRef] [PubMed]

- 15. Vickers, A.J. Decision analysis for the evaluation of diagnostic tests, prediction models and molecular markers. *Am. Stat.* **2008**, *62*, 314–320. [CrossRef] [PubMed]
- Ghatnekar, O.; Andersson, R.; Svensson, M.; Persson, U.; Ringdahl, U.; Zeilon, P.; Borrebaeck, C.A. Modelling the benefits of early diagnosis of pancreatic cancer using a biomarker signature. *Int. J. Cancer* 2013, 133, 2392–2397. [CrossRef] [PubMed]
- 17. Andersen, D.K.; Korc, M.; Petersen, G.M.; Eibl, G.; Li, D.; Rickels, M.R.; Chari, S.T.; Abbruzzese, J.L. Diabetes, Pancreatogenic Diabetes, and Pancreatic Cancer. *Diabetes* **2017**, *66*, 1103–1110. [CrossRef] [PubMed]
- Vasen, H.; Ibrahim, I.; Ponce, C.G.; Slater, E.P.; Matthai, E.; Carrato, A.; Earl, J.; Robbers, K.; van Mil, A.M.; Potjer, T.; et al. Benefit of Surveillance for Pancreatic Cancer in High-Risk Individuals: Outcome of Long-Term Prospective Follow-Up Studies From Three European Expert Centers. *J. Clin. Oncol.* 2016, *34*, 2010–2019. [CrossRef] [PubMed]
- 19. Norris, A.L.; Roberts, N.J.; Jones, S.; Wheelan, S.J.; Papadopoulos, N.; Vogelstein, B.; Kinzler, K.W.; Hruban, R.H.; Klein, A.P.; Eshleman, J.R. Familial and sporadic pancreatic cancer share the same molecular pathogenesis. *Fam. Cancer* **2015**, *14*, 95–103. [CrossRef] [PubMed]
- 20. Yue, T.; Partyka, K.; Maupin, K.A.; Hurley, M.; Andrews, P.; Kaul, K.; Moser, A.J.; Zeh, H.; Brand, R.E.; Haab, B.B. Identification of blood-protein carriers of the CA 19-9 antigen and characterization of prevalence in pancreatic diseases. *Proteomics* **2011**, *11*, 3665–3674. [CrossRef] [PubMed]
- 21. Capello, M.; Bantis, L.E.; Scelo, G.; Zhao, Y.; Li, P.; Dhillon, D.S.; Patel, N.J.; Kundnani, D.L.; Wang, H.; Abbruzzese, J.L.; et al. Sequential Validation of Blood-Based Protein Biomarker Candidates for Early-Stage Pancreatic Cancer. *J. Natl. Cancer Inst.* **2017**, *109*. [CrossRef] [PubMed]
- 22. Bunger, S.; Laubert, T.; Roblick, U.J.; Habermann, J.K. Serum biomarkers for improved diagnostic of pancreatic cancer: A current overview. *J. Cancer Res. Clin. Oncol.* 2011, 137, 375–389. [CrossRef] [PubMed]
- 23. Loosen, S.H.; Neumann, U.P.; Trautwein, C.; Roderburg, C.; Luedde, T. Current and future biomarkers for pancreatic adenocarcinoma. *Tumour Biol.* **2017**, *39*. [CrossRef] [PubMed]
- 24. Gallego, J.; Lopez, C.; Pazo-Cid, R.; Lopez-Rios, F.; Carrato, A. Biomarkers in pancreatic ductal adenocarcinoma. *Clin. Transl. Oncol.* **2017**, *19*, 1430–1437. [CrossRef] [PubMed]
- 25. Li, W.; Li, C.; Zhou, T.; Liu, X.; Liu, X.; Chen, D. Role of exosomal proteins in cancer diagnosis. *Mol. Cancer* 2017, *16*, 145. [CrossRef] [PubMed]
- Kaur, S.; Smith, L.M.; Patel, A.; Menning, M.; Watley, D.C.; Malik, S.S.; Krishn, S.R.; Mallya, K.; Aithal, A.; Sasson, A.R.; et al. A Combination of MUC5AC and CA19-9 Improves the Diagnosis of Pancreatic Cancer: A Multicenter Study. *Am. J. Gastroenterol.* 2017, 112, 172–183. [CrossRef] [PubMed]
- Kim, J.; Bamlet, W.R.; Oberg, A.L.; Chaffee, K.G.; Donahue, G.; Cao, X.J.; Chari, S.; Garcia, B.A.; Petersen, G.M.; Zaret, K.S. Detection of early pancreatic ductal adenocarcinoma with thrombospondin-2 and CA19-9 blood markers. *Sci. Transl. Med.* 2017, *9*, eaah5583. [CrossRef] [PubMed]
- 28. Poruk, K.E.; Firpo, M.A.; Scaife, C.L.; Adler, D.G.; Emerson, L.L.; Boucher, K.M.; Mulvihill, S.J. Serum osteopontin and tissue inhibitor of metalloproteinase 1 as diagnostic and prognostic biomarkers for pancreatic adenocarcinoma. *Pancreas* **2013**, *42*, 193–197. [CrossRef] [PubMed]
- 29. Radon, T.P.; Massat, N.J.; Jones, R.; Alrawashdeh, W.; Dumartin, L.; Ennis, D.; Duffy, S.W.; Kocher, H.M.; Pereira, S.P.; Guarner Posthumous, L.; et al. Identification of a Three-Biomarker Panel in Urine for Early Detection of Pancreatic Adenocarcinoma. *Clin. Cancer Res.* **2015**, *21*, 3512–3521. [CrossRef] [PubMed]
- 30. Faca, V.M.; Song, K.S.; Wang, H.; Zhang, Q.; Krasnoselsky, A.L.; Newcomb, L.F.; Plentz, R.R.; Gurumurthy, S.; Redston, M.S.; Pitteri, S.J.; et al. A mouse to human search for plasma proteome changes associated with pancreatic tumor development. *PLoS Med.* **2008**, *5*, e123. [CrossRef] [PubMed]
- 31. Iacobuzio-Donahue, C.A.; Ashfaq, R.; Maitra, A.; Adsay, N.V.; Shen-Ong, G.L.; Berg, K.; Hollingsworth, M.A.; Cameron, J.L.; Yeo, C.J.; Kern, S.E.; et al. Highly expressed genes in pancreatic ductal adenocarcinomas: A comprehensive characterization and comparison of the transcription profiles obtained from three major technologies. *Cancer Res.* **2003**, *63*, 8614–8622. [PubMed]
- 32. Debernardi, S.; Massat, N.J.; Radon, T.P.; Sangaralingam, A.; Banissi, A.; Ennis, D.P.; Dowe, T.; Chelala, C.; Pereira, S.P.; Kocher, H.M.; et al. Noninvasive urinary miRNA biomarkers for early detection of pancreatic adenocarcinoma. *Am. J. Cancer Res.* **2015**, *5*, 3455–3466. [PubMed]
- 33. Ioannidis, J.P. Why Most Clinical Research Is Not Useful. PLoS Med. 2016, 13, e1002049. [CrossRef] [PubMed]
- 34. Vidal, M. Interactome networks and human disease. *Cell* **2011**, 144, 986–998. [CrossRef] [PubMed]

- 35. Makohon-Moore, A.; Iacobuzio-Donahue, C.A. Pancreatic cancer biology and genetics from an evolutionary perspective. *Nat. Rev. Cancer* **2016**, *16*, 553–565.
- O'Brien, D.P.; Sandanayake, N.S.; Jenkinson, C.; Gentry-Maharaj, A.; Apostolidou, S.; Fourkala, E.O.; Camuzeaux, S.; Blyuss, O.; Gunu, R.; Dawnay, A.; et al. Serum CA19-9 is significantly upregulated up to 2 years before diagnosis with pancreatic cancer: Implications for early disease detection. *Clin. Cancer Res.* 2015, 21, 622–631. [CrossRef] [PubMed]
- 37. Hanash, S. A call for a fresh new look at the plasma proteome. *Proteom. Clin. Appl.* **2012**, *6*, 443–446. [CrossRef] [PubMed]
- Harsha, H.C.; Kandasamy, K.; Ranganathan, P.; Rani, S.; Ramabadran, S.; Gollapudi, S.; Balakrishnan, L.; Dwivedi, S.B.; Telikicherla, D.; Selvan, L.D.N.; et al. A compendium of potential biomarkers of pancreatic cancer. *PLoS Med.* 2009, *6*, e1000046. [CrossRef] [PubMed]
- 39. McShane, L.M. Statistical challenges in the development and evaluation of marker-based clinical tests. *BMC Med.* **2012**, *10*, 52. [CrossRef] [PubMed]
- Clarke, R.; Ressom, H.W.; Wang, A.; Xuan, J.; Liu, M.C.; Gehan, E.A.; Wang, Y. The properties of high-dimensional data spaces: Implications for exploring gene and protein expression data. *Nat. Rev. Cancer* 2008, *8*, 37–49. [CrossRef] [PubMed]
- 41. Kern, S.E. Why your new cancer biomarker may never work: Recurrent patterns and remarkable diversity in biomarker failures. *Cancer Res.* **2012**, *72*, 6097–6101. [CrossRef] [PubMed]
- 42. Ioannidis, J.P. Biomarker failures. Clin. Chem. 2013, 59, 202–204. [CrossRef] [PubMed]
- Mayo Clinic Prospective Resource for Biomarker Validation and Early Detection of Pancreatic Cancer. Mayo Clinic: Rochester, MN, USA, 2017. Available online: http://www.mayo.edu/research/centers-programs/ cancer-research/research-programs/gastrointestinal-cancer-program/mayo-clinic-pancreatic-cancerspore/core-resources/tissue-core (accessed on 29 November 2017).
- 44. Sakamoto, H.; Harada, S.; Nishioka, N.; Maeda, K.; Kurihara, T.; Sakamoto, T.; Higuchi, K.; Kitano, M.; Takeyama, Y.; Kogire, M.; et al. A Social Program for the Early Detection of Pancreatic Cancer: The Kishiwada Katsuragi Project. *Oncology* **2017**, *93*, 89–97. [CrossRef] [PubMed]
- 45. Lami, G.; Biagini, M.R.; Galli, A. Endoscopic ultrasonography for surveillance of individuals at high risk for pancreatic cancer. *World J. Gastrointest. Endosc.* **2014**, *6*, 272–285. [CrossRef] [PubMed]
- 46. Potjer, T.P.; Mertens, B.J.; Nicolardi, S.; van der Burgt, Y.E.; Bonsing, B.A.; Mesker, W.E.; Tollenaar, R.A.; Vasen, H.F. Application of a Serum Protein Signature for Pancreatic Cancer to Separate Cases from Controls in a Pancreatic Surveillance Cohort. *Transl. Oncol.* **2016**, *9*, 242–247. [CrossRef] [PubMed]
- 47. Ramsey, B.W.; Nepom, G.T.; Lonial, S. Academic, Foundation, and Industry Collaboration in Finding New Therapies. *N. Engl. J. Med.* **2017**, *376*, 1762–1769. [CrossRef] [PubMed]
- 48. Stern, A.D.; Alexander, B.M.; Chandra, A. Innovation Incentives and Biomarkers. *Clin. Pharmacol. Ther.* **2018**, 103, 34–36. [CrossRef] [PubMed]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).