OPEN ACCESS

Original research

Development and validation of a serum microRNA biomarker panel for detecting gastric cancer in a high-risk population

Jimmy Bok Yan So,[1,2,3] Ritika Kapoor,[4] Feng Zhu ,[5] Calvin Koh ,[6] Lihan Zhou,[7] Ruiyang Zou,[7] Yew Chung Tang,[7] Patrick C K Goo,[8] Sun Young Rha,[9] Hyun Cheol Chung,[9] Joanne Yoong,[10] Celestial T Yap,[11] Jaideepraj Rao,[12] Chung-King Chia,[13] Stephen Tsao,[13] Asim Shabbir,[3] Jonathan Lee ,[5,6] Kong-Peng Lam ,[14] Mikael Hartman,[1,10] Wei Peng Yong,[15] Heng-Phon Too,[16] Khay-Guan Yeoh [5,6,17]

► Additional material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2020-322065).

For numbered affiliations see end of article.

Correspondence to
Dr Khay-Guan Yeoh, Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, 119228, Singapore; mdcykg@nus.edu.sg

Received 2 June 2020
Revised 8 September 2020
Accepted 9 September 2020
Published Online First 7 October 2020

Watch Video
gut.bmj.com

To cite: So JBY, Kapoor R, Zhu F, *et al*. *Gut* 2021;**70**:829–837.

ABSTRACT

Objective An unmet need exists for a non-invasive biomarker assay to aid gastric cancer diagnosis. We aimed to develop a serum microRNA (miRNA) panel for identifying patients with all stages of gastric cancer from a high-risk population.

Design We conducted a three-phase, multicentre study comprising 5248 subjects from Singapore and Korea. Biomarker discovery and verification phases were done through comprehensive serum miRNA profiling and multivariant analysis of 578 miRNA candidates in retrospective cohorts of 682 subjects. A clinical assay was developed and validated in a prospective cohort of 4566 symptomatic subjects who underwent endoscopy. Assay performance was confirmed with histological diagnosis and compared with *Helicobacter pylori* (HP) serology, serum pepsinogens (PGs), 'ABC' method, carcinoembryonic antigen (CEA) and cancer antigen 19–9 (CA19-9). Cost-effectiveness was analysed using a Markov decision model.

Results We developed a clinical assay for detection of gastric cancer based on a 12-miRNA biomarker panel. The 12-miRNA panel had area under the curve (AUC)=0.93 (95% CI 0.90 to 0.95) and AUC=0.92 (95% CI 0.88 to 0.96) in the discovery and verification cohorts, respectively. In the prospective study, overall sensitivity was 87.0% (95% CI 79.4% to 92.5%) at specificity of 68.4% (95% CI 67.0% to 69.8%). AUC was 0.848 (95% CI 0.81 to 0.88), higher than HP serology (0.635), PG 1/2 ratio (0.641), PG index (0.576), ABC method (0.647), CEA (0.576) and CA19-9 (0.595). The number needed to screen is 489 annually. It is cost-effective for mass screening relative to current practice (incremental cost-effectiveness ratio=US$44 531/quality-of-life year).

Conclusion We developed and validated a serum 12-miRNA biomarker assay, which may be a cost-effective risk assessment for gastric cancer.

Trial registration number This study is registered with ClinicalTrials.gov (Registration number: NCT04329299).

Significance of this study

What is already known on this subject?

► Circulating microRNAs (miRNAs) are promising biomarkers for detection of gastric cancer (GC) but previous studies were limited by small cohort sizes and the use of research grade assays.

What are the new findings?

► This is the most extensive evaluation of circulating miRNAs as biomarkers for GC detection to date, as measured by both cohort size and technical stringency. A clinical assay based on a panel of 12-miRNA biomarkers was developed, manufactured to clinical standards, and prospectively validated in a multicentre cohort of over 5000 subjects. The assay was more accurate than any existing blood-based diagnostic biomarkers for GC and it could reduce the number of unnecessary upper endoscopy. It was also cost-effective as a screening test for GC.

How might it impact on clinical practice in the foreseeable future?

► The serum miRNA biomarker panel can be used as a risk assessment tool for GC before endoscopy. It has the potential to be a cost-effective mass screening tool for GC.

INTRODUCTION

Gastric cancer (GC) is the fifth most common cancer and the third-leading cause of cancer deaths worldwide.[1] GC mortality is high due to late presentation.[2] In high-incidence countries, such as Japan and Korea, mass screening for GC is practiced using photofluorography or, more recently, endoscopy. In these settings, over 50% of GC patients are diagnosed at early stages and their survival is excellent.[3 4] However, in most countries, mass screening is neither feasible nor cost-effective because such screening methods are costly and invasive, with

poor compliance.[4] There is an unmet need for a less invasive and cost-effective GC screening test.

Currently available gastrointestinal tumour markers, including carcinoembryonic antigen (CEA) and cancer antigen 19–9 (CA19-9) are inadequate for GC screening due to their poor sensitivities, especially for early-stage GC.[5] Recently, the ABC method, a combination of age, serum anti-*Helicobacter pylori* (HP) IgG antibody (serology) and pepsinogen (PG) 1 and 2 levels, have shown some promise as a blood test for GC risk stratification in Japan but its clinical performance has varied among different populations.[4 6 7]

MicroRNAs (miRNAs or miRs) are small non-coding RNAs that regulate gene expression post-transcriptionally.[8] Aberrant expression of miRNAs has been implicated in the pathogenesis of many diseases, including cancer.[9 10] Cell-free miRNAs have been shown to circulate stably in serum and plasma,[11 12] and dysregulation of their expressions correlate with cancer onset and progression, making them attractive biomarker candidates.[12 13] However, sensitive and robust detection of these circulating miRNA from clinical samples is challenging due to their small size and low abundance.[14] To overcome these challenges, we developed a proprietary miRNA RT-qPCR assay platform with greater sensitivity and reproducibility in detecting circulating miRNAs using small-volume clinical samples.[15]

The primary aim of this study is to develop a serum panel of miRNA as a non-invasive test that can detect GC of all stages and validate its performance in a large prospective cohort. The secondary aim is to evaluate its cost-effectiveness as a mass screening tool for GC.

METHODS

Study design and patient cohorts

We conducted a three-phase, multicentre study to discover and validate a panel of serum miRNA biomarkers for GC. In the discovery phase, we measured the expression of 578 circulating miRNAs in a case–control cohort of 472 Singaporean Chinese subjects, including 236 cancer and 236 matched control subjects, to identify candidate biomarker miRNAs as well as candidate multi-miRNA panels. A total of 236 patients with cancer were from the Gastric Cancer Biomarker Discovery Study (GASCAD), which recruited newly diagnosed GC patients. Blood was collected prior to any cancer treatment. Matched control subjects were enrolled through the Gastric Epidemiology and Molecular Genetics Project (GCEP), a prospective cohort study that aimed to identify GC risk factors in the Singapore Chinese population with age 50 or above and to develop a screening algorithm.[16] All control subjects received surveillance endoscopy with standardised biopsy protocol at regular intervals and were confirmed to have no GC or high-grade dysplasia based on endoscopy and histological examination. GC patients and controls were matched in ethnicity (Chinese), sex and age (±10 years).

In the verification phase, we confirmed the dysregulation of candidate biomarkers and identified a 12-miR panel in another case–control cohort of 210 Singaporean and Korean subjects, including 94 cancers and 116 matched controls. Blinded biomarker verification was performed with sera specimen from cancer and control subjects recruited from Singapore and Korea. The Singaporean sample set included 20 additional GC patients and 69 matched controls from GASCAD and GCEP cohorts respectively. The Korean sample set included 74 GC patients recruited at Yonsei Cancer Center and 47 controls who were healthy blood donors from Songdang Institute for Cancer Research.

After the verification phase, a clinical grade multivariate index assay based on the 12-miR panel was formulated. Finally, the performance of this 12-miR panel was validated in a prospective cohort of 4566 Singaporean subjects who underwent upper endoscopy for their gastrointestinal symptoms. Patients eligible for inclusions were consecutive adults, between the ages of 40 and 90 years, who were scheduled to undergo gastroscopy based on standard clinical indications at National University Hospital and Tan Tock Seng Hospital in Singapore from 2013 to 2016. A total of 5282 subjects were recruited. Subjects with a history of total or partial gastrectomy were excluded. The presence and absence of GC and high-grade dysplasia were confirmed by endoscopy and histological examinations. Written informed consent was obtained from all participants.

Blood collection and serum processing

Fasting blood samples (20 mL) were withdrawn from each subject via venipuncture and collected in two plain serum tubes (BD vacutainer plus plastic serum tube). The serum tubes were centrifuged at 3000 rpm at 20°C for 10 min. Centrifugation and serum collection was done within 4 hours of blood collection. Serum specimen were aliquoted and stored immediately at −80°C.

MiRNA quantification in discovery and verification phases

The absolute expression (copy numbers) of 578 candidate miRNAs were quantified in each patient and control biospecimen using miRNA-specific RT-qPCR assays (MiRXES, Singapore) via a highly controlled workflow illustrated in online supplemental figure S1. The analytical specificity, reproducibility and sensitivity of the assay and workflow (online supplemental methods) are shown in online supplemental figure S2. Total RNA from 200 µL of patient and control serum specimen was isolated using miRNeasy serum/plasma miRNA isolation kit (Qiagen, Germany). Synthetic miRNA controls were added to samples before RNA isolation, and RT-qPCR to monitor and normalise technical variations throughout the entire workflow (online supplemental methods). Absolute expression of each miRNA was determined in each patient serum sample and normalised across samples using endogenous reference miRNAs (online supplemental methods).

12-miR assay in validation phase

A central biorepository received all serum specimens. The 12-miR qPCR assay was developed and manufactured in accordance with the ISO13485 medical devices quality management systems (MIRXES, Singapore). Laboratory testing (online supplemental methods) was performed in CAP/ISO-certified laboratories without knowledge of the results of endoscopic and histopathological findings. Eleven GC related miRNAs (miR-140, miR-183, miR-30e, miR-103a, miR-126, miR-93, miR-142, miR-21, miR-29c, miR-424 and miR-181a) were measured together with a reference miRNA (miR-340) (online supplemental methods). All qPCR measurements were performed in duplicates. The assay generated a numerical GC risk score for each sample using the GASTROSmart Software (MIRXES, Singapore) (online supplemental methods). Using the 12-miRNA panel, a cancer prediction score was generated based on the most optimal sensitivity and specificity combination. The risk score was calculated using a linear regression model of the measured expression levels of the 12 miRNAs in the panel. A score of 40 or more was defined as a positive test result.

Other blood-based GC marker assays

Serum biomarkers CEA, CA19-9, anti-HP IgG, PG 1 and PG 2 were also measured. CEA, CA19-9, PG 1/2 ratio, PG index

(combining a PG 1 level cut-off and PG 1/2 ratio), and the so-called 'ABC method' by combining HP serology and PG 1/2 ratio have previously been proposed as biomarkers for GC screening.[4 6 7 17] Serum specimens were qualitatively assayed for HP antibodies with the MP Diagnostics HELICO BLOT 2.1 Western blot kit assay (MP BIOMEDICALS Asia Pacific). PG I and II levels were determined with latex agglutination turbidimetric immunoassay kit (LZ Test 'Eiken' PG I and II, Tokyo, Japan). Access CEA and CA19-9 chemiluminescent sandwich immunoassays (Beckman Coulter, USA) were run on the UniCel DxI 800 immunoassay system (Beckman Coulter, USA). All assays were performed without knowledge of the clinical findings.

Health economic analysis

The overall costs and health benefits of conducting mass screening using the validated miRNA panel relative to current practice of no-screening were estimated for a hypothetical cohort, with assumed health seeking behaviours, disease incidence/progression and associated patient quality-of-life years (QALY) representative of a high-risk population in Singapore (Chinese males of 50–75 years old). We assumed that miRNA test-positive subjects would go on to have a confirmatory endoscopy while test-negative subjects would be followed up in 3 years. Further modelling and parameter uncertainties were addressed using sensitivity analysis. Details of this analysis are provided in online supplemental methods.

Statistical analysis

During the discovery phrase, significantly regulated miRNAs were identified using Student's t-tests with a false discovery rate (FDR) correction.[18] The receiver operating characteristic curve was used to present the performance of individual miRNA or multivariant biomarker panels. A sequential forward floating search (SFFS) algorithm[19] was used to optimise the miRNA biomarker panel, with area under the curve (AUC) values as the performance indicator. Multivariate analysis was carried out to construct multi-miR panels with associated algorithms for classifying cancers and controls (online supplemental methods). Multivariate Cox regression analysis of the 12-miR panel and clinical covariates was carried out.

For the prospective validation study, the sample size was calculated based on an expected 2% prevalence of GC in a high-risk symptomatic population and a point estimate of 85% sensitivity based on discovery and verification data. We planned to recruit about 5000 participants to achieve margins of sampling error of approximately 5 percentage points for sensitivity.

The study was reported according to Standards for Reporting of Diagnostic Accuracy Studies 2015 guidelines.[20]

RESULTS

Identification of GC associated serum miRNA biomarkers

GC-associated serum miRNA biomarkers were identified through retrospective analysis of 472 prospectively collected specimens from GC patients and controls matched by age, sex and ethnicity. Clinicopathological characteristics of the discovery cohort are shown in table 1. The cohort was enriched for early stage GC patients (30.1% stage 1 and 15.3% stage 2) to ensure that it was sufficiently powered to identify biomarkers associated with early stage GC. We systematically evaluated an a priori list of 578 circulating miRNAs using a highly controlled and analytically validated RT-qPCR workflow. Prior to biomarker discovery, the analytical sensitivity, specificity and reproducibility of the miRNA assay and workflow were validated using a combination of synthetic miRNA templates. These assays demonstrated strong discrimination against highly homologous miRNA sequences (online supplemental figure S2A), high concordance in detecting circulating miRNAs in both control and cancer sera (online supplemental figure S2B), and good dynamic range in amplifying and detecting miRNAs with distinct sequences and varying AT content (online supplemental figure S2C).

Of 578 serum miRNAs quantified, 191 miRNAs were detected in more than 90% of the subjects (expression levels ≥500 copies/mL of serum) (online supplemental table S1); 75 of the 191 miRNAs were differentially expressed between cancer patients and matched controls (FDR corrected $p<0.01$) (online supplemental table S2). Of the 75 dysregulated miRNAs, 68 were novel discoveries; 51 were upregulated and 24 downregulated (online supplemental tables S2 and S3). An expression heatmap of the dysregulated miRNAs showed the majority of GC subjects clustered closely (figure 1A). Many of these miRNAs were positively correlated (Pearson correlation coefficient (r) >0.6) (figure 1B), suggesting potential co-regulation in miRNA expression. Among the dysregulated miRNAs, miR-142–5 p (upregulated) and miR-99b-5p (downregulated) exhibited the highest AUC, at 0.71 and 0.67, respectively (online supplemental figure S3A). Seven miRNAs were found to be differentially expressed (FDR corrected $p<0.05$) in the three histological subtypes (diffuse, intestinal and mixed) of GC (online supplemental figure S4). Thirty-six miRNAs were differentially regulated at various stages of GC (FDR corrected $p<0.01$) (online supplemental figure S3B and table S4).

Table 1 Discovery cohort clinicopathological characteristics

Singaporean			
Discovery phase		**Case–control cohort**	
Subjects cohort size		**Control subjects n=236**	**Patients with cancer n=236**
Gender	Male	150 (63.3%)	148 (62.7%)
	Female	87 (36.7%)	88 (37.3%)
Age		61.2±8.4 (SD)	68.0±10.9 (SD)
Ethnicity	Chinese (%)	236 (100%)	236 (100%)
Stages (AJCC 2010)	Stage 0 (%)	–	–
	Stage 1 (%)	–	71 (30.1%)
	Stage 2 (%)	–	36 (15.3%)
	Stage 3 (%)	–	54 (22.9%)
	Stage 4 (%)	–	75 (31.8%)
	Unknown (%)	–	–
Histological subtype	Intestinal	–	134 (56.8%)
	Diffuse	–	70 (29.7%)
	Mixed	–	32 (13.6%)
	Unknown	–	–
Gastritis	No	7 (3.0%)	36 (15.3%)
	Yes	230 (97.0%)	200 (84.7%)
	Unknown	–	–
Intestinal metaplasia	No	116 (48.9%)	75 (31.8%)
	Yes	121 (51.1%)	161 (68.2%)
	Unknown	–	–
Atrophy	No	133 (56.1%)	215 (91.1%)
	Yes	104 (43.9%)	21 (8.9%)
	Unknown	–	–
Helicobacter pylori	No	105 (44.3%)	50 (21.2%)
	Yes	132 (55.7%)	186 (78.8%)

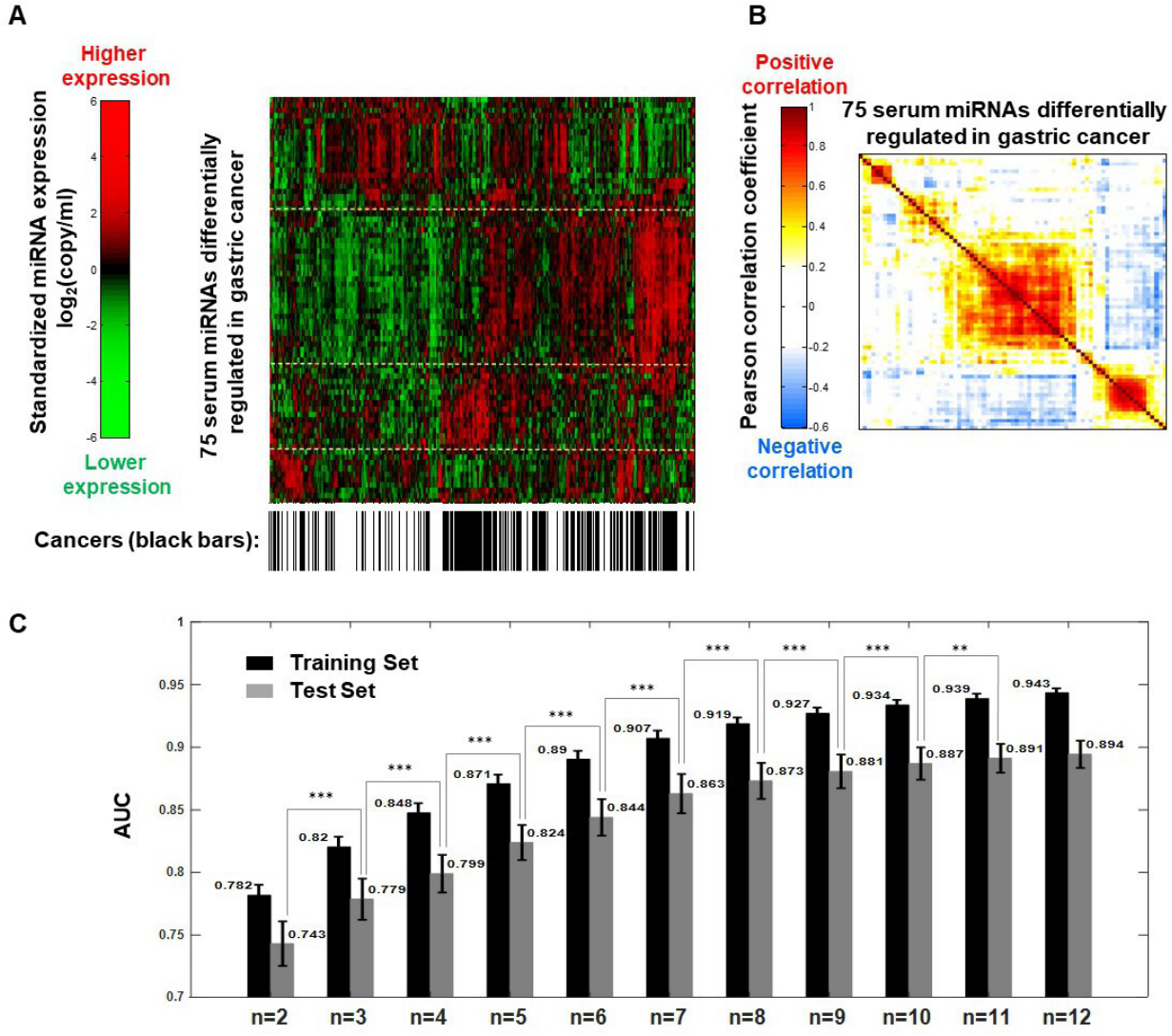


Figure 1 Identification of candidate miRNA biomarkers and multi-miRNA biomarker panels for gastric cancer detection. (A) Heat-map showing expression levels of serum miRNAs that were differentially regulated in gastric cancer. The full list can be found in in online supplemental table S2; absolute miRNA expression levels (copy/mL) of miRNAs were presented in log2 scale and standardised to zero mean. Hierarchical clustering was carried out for both dimensions (miRNAs and samples) based on Euclidean distance. (B) Correlation in expression levels between differentially regulated miRNAs. Pearson's linear correlation coefficients were calculated between all 75 miRNAs that were identified to be differentially regulated in gastric cancer (online supplemental table S2). (C) Gastric cancer detection accuracy of multi-miRNA biomarker panels with 3–10 miRNAs as determined by mean area under ROC curve (AUC). Biomarker panels were tested in the discovery cohort. Two hundred iterations of a cross-validation process were carried out by dividing the Discovery cohort into two data sets: training and testing. Error bars indicate SD. Statistical significance of difference in AUC was determined using Student's t-test (one sided, **p<0.01; ***p<0.001). AUC, area under the curve; miRNA, micro-RNA.

Next, we developed and tested multi-miRNA biomarker panels with high AUC in distinguishing cancer from controls using cross-validation. The discovery cohort was partitioned into equally sized training and test sets matched for cancer stage, subtype, age, gender and ethnicity.[21] We derived multi-miRNA biomarker panels in training set using SFFS and SVM and tested the panel performance in the test set. The composition of the miRNA panels included a combination of miRNAs that were individually significant and non-significant between GC and matched controls. An improvement in AUC was observed when number of miRNAs in the panel increased but plateaued at 12-miRNAs. The median AUC values for a 12-miRNA panel were close to 0.90 in the test set, with a spread between the 25th and 75th percentile of <0.05 (figure 1C). Incorporating more miRNAs into the panel did not significantly improve AUC. Multivariate Cox regression analysis showed that the 12-miRNA biomarker panel was independent of clinical covariates in detecting GC (online supplemental table S5).

Verification of miRNA biomarker panel

We verified the performance of individual miRNAs and the 12-miR panel in two independent retrospective case–control cohorts of 89 subjects from Singapore and 121 subjects from Korea (table 2). We observed good correlations in individual miRNA expression fold-changes between the discovery cohort and each of the two verification cohorts (figure 2A).

Similarly, the 12-miR panel identified through the discovery cohort showed consistency in the verification cohorts. The panel was able to discriminate GC from matched controls with AUC 0.93

Table 2 Verification cohort clinicopathological characteristics

Verification phase		Singaporean Case–control cohort		Korean Case–control cohort	
Subjects cohort size		Control subjects n=69	Patients with cancer n=20	Control subjects n=47	Patients with cancer n=74
Gender	Male	32 (46.4%)	14 (70.0%)	35 (74.5%)	44 (59.5%)
	Female	37 (53.6%)	6 (30.0%)	12 (25.5%)	30 (40.5%)
Age		63.3±8.4 (SD)	74.8±9.5 (SD)	26.4±2.7 (SD)	59.1±10.6 (SD)
Ethnicity	Chinese (%)	69 (100%)	20 (100%)	–	–
	Korean (%)	–	–	47 (100%)	74 (100%)
Stages (AJCC 2010)	Stage 0 (%)	–	–	–	–
	Stage 1 (%)	–	10 (50.0%)	–	17 (23.0%)
	Stage 2 (%)	–	3 (15.0%)	–	21 (28.4%)
	Stage 3 (%)	–	6 (30.0%)	–	17 (23.0%)
	Stage 4 (%)	–	1 (5.0%)	–	19 (25.7%)
	Unknown (%)	–	–	–	–
Histological subtype	Intestinal	–	–	–	35 (47.3%)
	Diffuse	–	–	–	31 (41.9%)
	Mixed	–	–	–	0 (0.0%)
	Unknown	–	–	–	8 (10.8%)
Gastritis	No	0 (0%)	0 (0%)	–	–
	Yes	69 (100%)	12 (60%)	–	–
	Unknown	–	8 (40%)	–	–
Intestinal metaplasia	No	17 (24.6%)	0 (0.0%)	–	–
	Yes	52 (75.4%)	10 (50.0%)	–	–
	Unknown	–	10 (50.0%)	–	–
Atrophy	No	32 (46.4%)	0 (0.0%)	–	–
	Yes	37 (53.6%)	1 (5.0%)	–	–
	Unknown	–	19 (95.0%)	–	–
H Pylori	No	18 (26.1%)	2 (10.0%)	–	–
	Yes	51 (73.9%)	18 (90.0%)	–	–

(95% CI 0.90 to 0.95) in discovery cohort and 0.92 (95% CI 0.88 to 0.96) in the verification cohort. When comparing early-stage GC (stages 1–2) from matched controls, the panel achieved AUC 0.90 (95% CI 0.85 to 0.94) in the discovery cohort and 0.91 (95% CI 0.85 to 0.96) in the verification cohort (figure 2B). The verified 12-miR panel (miR-140, miR-183, miR-30e, miR-103a, miR-126, miR-93, miR-142, miR-21, miR-29c, miR-424, miR-181a and miR-340) was thus finalised and developed into a clinical assay in accordance with the ISO13485 medical devices quality management system for prospective validation.

Prospective validation of the 12-miR assay

Study population

The 12-miR assay, with a prespecified prediction algorithm, was validated in a large prospective validation cohort of Singaporean patients (figure 3). The clinicopathological characteristics are shown in table 3. A total of 5282 participants underwent endoscopy, and had serum collected for testing with the 12-miR assay and other serum-based biomarker tests that have been suggested for GC detection (HP serology, PG 1/2 ratio, PG index, ABC method, CEA and CA19-9). 597 subjects were excluded from miRNA analysis due to sample quality issues. Of the remaining 4685 samples assayed, 4570 (97.5%) yielded valid test results. The remaining samples did not yield valid results due to invalid expression ranges or were excluded from data analysis due to incomplete clinical information. Altogether, 4566 participants had results that could be fully analysed for the 12-miR assay, HP serology, CEA and CA19-9 tests. A total of 133 samples were excluded from PG 1/2 analysis because patients had renal failure which affects PG levels. A total of 4433 patients had results for the PG tests which could be analysed. A total of 115 biopsy-proven GC was found by endoscopy (prevalence, 2.5%). Another 10 participants were found to have gastric high-grade dysplasia.

Assay performance

The 12-miR distinguished GC from matched normal controls with AUC 0.848 (95% CI 0.809 to 0.880) (figure 4A). This GC detection accuracy was higher than HP serology (AUC 0.635, 95% CI 0.594 to 0.668), PG 1/2 ratio (AUC 0.641, 95% CI 0.567 to 0.705), PG index (AUC 0.576, 95% CI 0.540 to 0.626), ABC method (AUC 0.647, 95% CI 0.60 to 0.681), CEA (AUC 0.576, 95% CI 0.512 to 0.638) or CA19-9 (AUC 0.595, 95% CI 0.535 to 0.656) (figure 4A,B).

The 12-miR assay identified 100 of 115 GC detected by endoscopy, for an overall sensitivity of 87.0% (95% CI 79.4% to 92.5%) at a specificity of 68.4% (95% CI 67.0% to 69.8%). The 12-miR assay achieved the highest sensitivity among the serum-based biomarker tests (figure 4C). PG, CEA and CA19-9 tests had higher than 90% specificity but lower than 30% sensitivity for GC detection. GC detection accuracy with the 12-miR assay can be improved by including patient's age, HP serology and PG 1/2 ratio (figure 4D). Using this combination, AUC was improved to 0.884, with specificity of 69.4% at sensitivity of 87.0%. GC detection sensitivity of the 12-miR did not vary significantly by

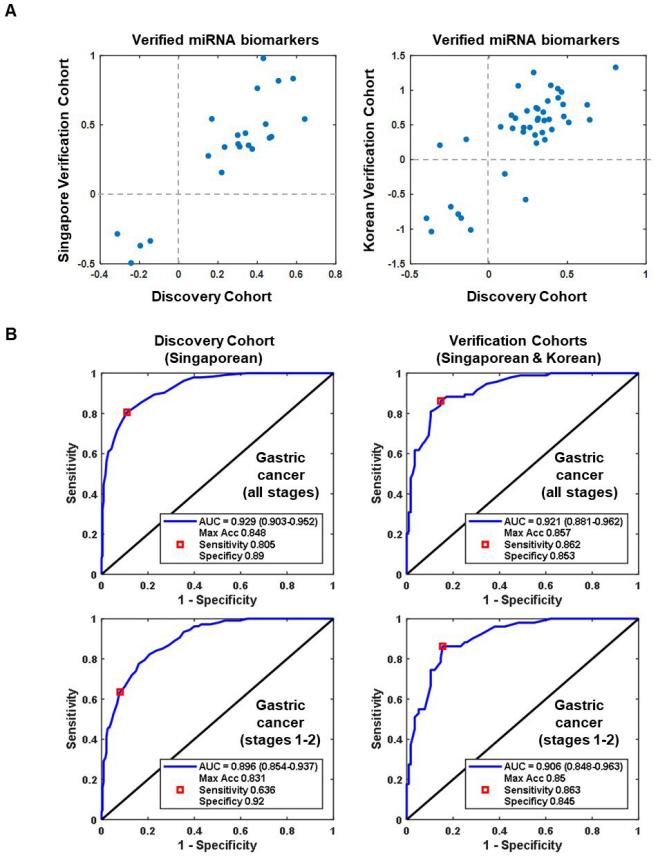


Figure 2 Verification of gastric cancer miRNA biomarkers and multi-miRNA biomarker panel detection accuracy in independent cohorts. (A) Correlation in expression level fold changes (cancer over control) of verified miRNA biomarkers between the discovery cohort and verification cohorts. (B) Receiver operating characteristics (ROC) curves for the 12-miRNA biomarker panel in detecting all gastric cancers (A) and early stage (stage 1–2) cancers (B). Area under the ROC curve (AUC) used to determine gastric cancer detection accuracy. Maximum classification accuracy is determined to occur at the point indicated by the red box. The sensitivity and specificity at this point is shown. miRNA, micro-RNA.

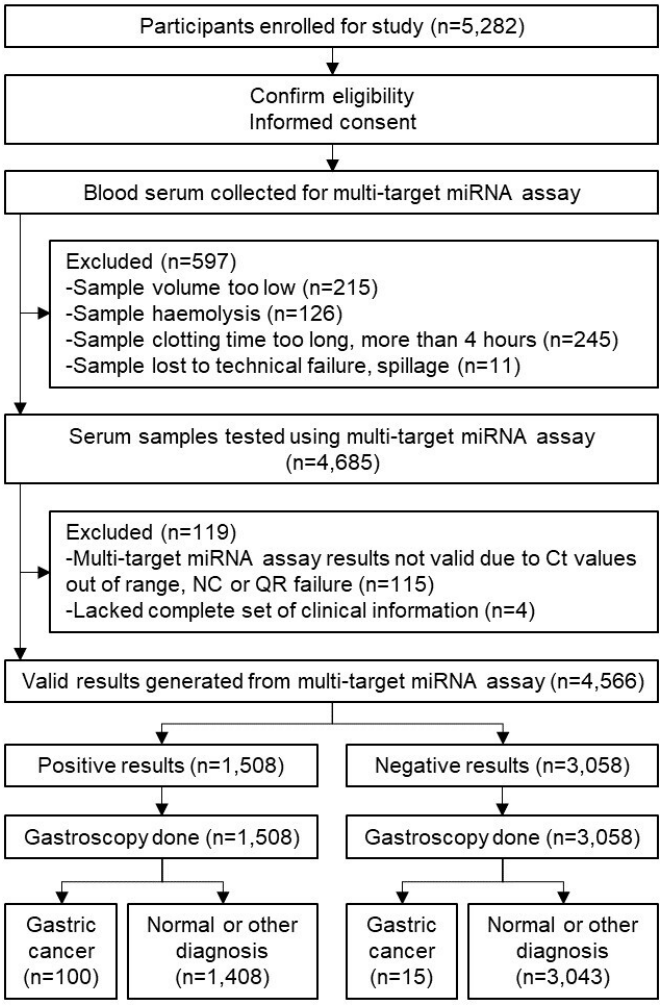


Figure 3 Prospective validation of 12-miR biomarker assay for detection of gastric cancer. Flow chart of prospective validation study design prepared in accordance with Standards for Reporting of Diagnostic Accuracy Studies guidelines. miR, micro RNA, NC, negative control; QR, quantitative reference.

cancer stage, gender and ethnicity but tends to be higher in older patients, larger tumour and intestinal-type GCs (figure 5).

To detect 115 GC cases in a symptomatic population, 4566 gastroscopies were carried out. Therefore, 40 gastroscopies will be required to detect one case of GC if no biomarker test were used. In comparison, only 15 endoscopies will be required if the 12-miR assay result was used to select patients for endoscopy in the same population. The positive predictive value (PPV) of using the 12-miR assay is 6.7% while the negative predictive value (NPV) is 99.5%. The assay had minimal cross-reactivity with other common cancers including those of the gastrointestinal tract (online supplemental table S6 and figure S6).

Health economic analysis

Using an assumption of compliance reflective of the existing screening programmes in Asia (45%), the number needed to screen (NNS) with the 12-miR assay in order to detect one case of GC was calculated to be 489 annually. Mass screening with the 12-miR assay can increase early-stage GC detection rate to 40 per 100 cancers identified, compared with 30 per 100 in current practice (table 4). Overall, mass screening using the 12-miR assay added costs of USD 175 per subject. The ratio of additional cost to additional health gains, or the incremental cost-effectiveness ratio is USD44 531/QALY, which is cost-effective in the local context, compared with the WHO-CHOICE threshold of approximately USD50 000/QALY (table 5). We anticipate that actual compliance for mass screening is likely to be significantly higher for a non-invasive test compared with endoscopy.

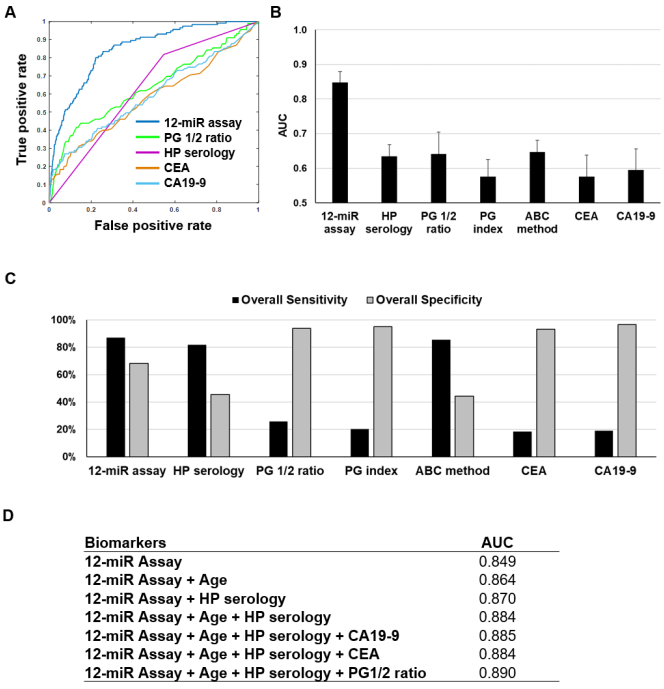
DISCUSSION

In recent years, miRNAs have been investigated as promising GC biomarkers because many solid tumours exhibit dysregulation of miRNA expression.[13] Patients with cancer exhibit aberrant expression of circulating miRNAs in biofluids such as blood.[22 23] However, change in miRNA expression in blood is less readily detected than changes in tissue due to multiple tissue sources for circulating miRNA and multiple physiological or pathological conditions affecting miRNA quantities. It is also technically challenging to detect miRNAs because of their small size. Previous studies exploring circulating miRNAs as GC biomarkers have shown promising proof-of-concept results but remain largely inconclusive, possibly due to small cohort sizes (n=6 to 570),[24–26] and the use of research grade assays.[27]

Table 3 Prospective validation cohort clinicopathological characteristics

Singaporean			
Validation phase		**Prospective cohort**	
		Control subjects	**Patients with cancer**
Subjects cohort size		**n=4441**	n=125
Gender	Male	2346 (52.83%)	76 (60.80%)
	Female	2095 (47.17%)	49 (39.20%)
Age		57.17±10.48 (SD)	56.90±10.31 (SD)
Ethnicity	Chinese (%)	3394 (76.42%)	96 (76.80%)
	Malay (%)	325 (7.32%)	7 (5.60%)
	Indian (%)	369 (8.31%)	9 (7.20%)
	Others (%)	353 (7.95%)	13 (10.40%)
Stages (AJCC 2010)	Stage 0 (%)	–	10 (8.00%)
	Stage 1 (%)	–	16 (12.80%)
	Stage 2 (%)	–	20 (16.00%)
	Stage 3 (%)	–	31 (24.80%)
	Stage 4 (%)	–	38 (30.40%)
	Unknown (%)	–	10 (8.00%)
Histological subtype	Intestinal	–	39 (31.2%)
	Diffuse	–	30 (24%)
	Mixed	–	13 (10.4%)
	Unknown	–	43 (34.4%)
Intestinal metaplasia	No	1936 (43.6%)	29 (23.2%)
	Yes	609 (13.7%)	55 (44%)
	Unknown	1896 (42.7%)	41 (32.8%)
Atrophy	No	2505 (56.4%)	42 (33.6%)
	Yes	37 (0.833%)	5 (4%)
	Unknown	1899 (42.8%)	78 (62.4%)
Helicobacter pylori	No	2015 (45.37%)	21 (16.80%)
	Yes	2426 (54.63%)	104 (83.20%)

AJCC, American Joint Committee on Cancer.



D

Biomarkers	AUC
12-miR Assay	0.849
12-miR Assay + Age	0.864
12-miR Assay + HP serology	0.870
12-miR Assay + Age + HP serology	0.884
12-miR Assay + Age + HP serology + CA19-9	0.885
12-miR Assay + Age + HP serology + CEA	0.884
12-miR Assay + Age + HP serology + PG1/2 ratio	0.890

Figure 4 Gastric cancer detection accuracy of 5-miR biomarker assay compared with other serum-based biomarker tests. (A) ROC curves for 12-miR assay, PG 1/2 ratio, HP serology, CEA, and CA19-9 for detection of gastric cancer. (B) AUC for 12-miR biomarker assay compared with HP serology, PG 1/2 ratio, PG index, ABC method, CEA, and CA19-9 tests. Bars show 95% CI (C) Overall sensitivity and associated specificity of GC detection using the 12-miR assay (both high sensitivity and high specificity cut-offs), HP serology, PG 1/2 ratio, PG index, ABC method, CEA, and CA19-9 tests. (D) Combinations of biomarker tests with optimal AUC for detecting gastric cancer. AUC, area under the curve; CA19, cancer antigen 19; CEA, carcinoembryonic antigen; HP, Helicobacter pylori; miR, micro-RNA; PG, pepsinogen.

To the best of our knowledge, this is the most extensive evaluation of circulating miRNAs as biomarkers for GC detection as measured by both cohort size and technical stringency. We quantified the absolute expression of 578 serum miRNAs in 682 cancer patients and control subjects using a comprehensive biomarker detection platform that incorporated important advances that were designed to overcome the biological and technical challenges inherent in detecting circulating miRNA. Sixty-eight novel serum miRNAs associated with GC were discovered. Subsequently, we used multivariant data analysis to develop a 12-miR assay that discriminated between GC patients and matched high-risk controls with high accuracy (AUC >0.93). A clinical grade 12-miR assay was then manufactured in accordance with the ISO13485 medical devices quality management systems and validated in a prospective cohort of 4566 patients. Depending on cut-off used, detection sensitivity reached 87% while specificity was as high as 93.9%. The assay could detect GC with high sensitivity across all age groups, genders, ethnicities and tumour stages. The serum 12-miR assay performed significantly better than any of the conventional blood-based biomarker tests. Its performance can be further enhanced by combining with age and HP serology to achieve an AUC of 0.884. Furthermore, we demonstrated the clinical specificity of the 12-miR assay against seven other prevalent cancers, including lung, breast, colorectal, liver, oesophageal, prostate and bladder cancer. We used a multivariate panel, instead of solitary miRNA biomarkers, to overcome low detection accuracy attributed to tumour heterogeneity.[12]

In the symptomatic study population, the 12-miR assay has a NNS to detect one cancer of 15, comparing favourably with NNS of 40 with unselected gastroscopy in the same population. While it is not the intent of the 12-miR assay to replace endoscopic evaluation, we believe this assay provides a useful option for symptomatic patients who might not be keen on initial endoscopic screening. The 12-miR assay is also a potential tool for mass screening. In this scenario, NNS was 489 annually which compares favourably with prostate-specific antigen, a common serum-based screening test for prostate cancer.[28] If used for mass screening, the 12-miR assay is also capable of detecting more early-stage GCs than current practice. Furthermore, a blood-based test is expected to have better population compliance compared with a scope-based evaluation. In countries with high prevalence of GC, and with endoscopic screening programmes, such as South Korea, the 12-miR assay can enhance overall compliance among the population subset who decline endoscopy. In countries with intermediate prevalence of GC and no current screening programmes, such as Singapore, the 12-miR assay can be implemented as a screening test with endoscopic examinations for those with positive 12-miR assay results.

There are limitations of this assay. First, the 12-miR panel had high sensitivity but moderate specificity. We chose a risk score cut-off giving high sensitivity by design to minimise false

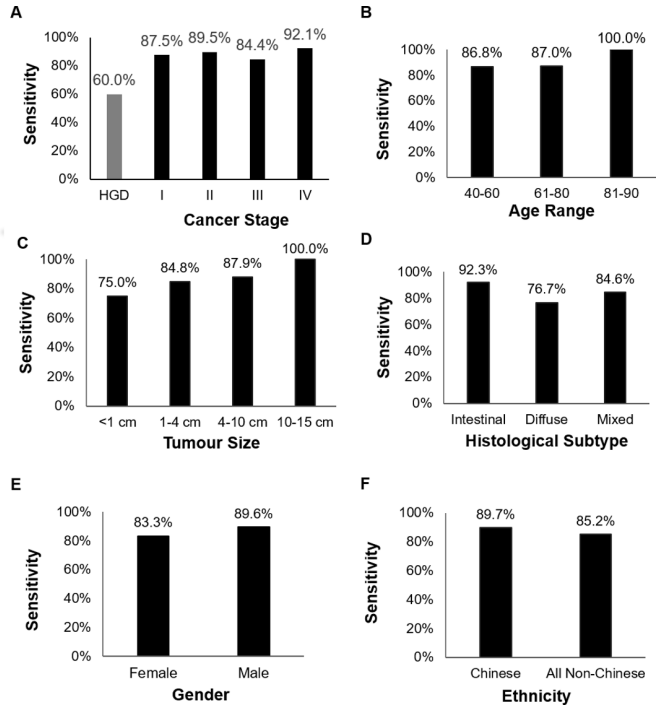


Figure 5 Detection sensitivity of 12-miR assay by gastric cancer stage and clinicopathological characteristics. Detection sensitivity at 68.4% specificity according to (A) gastric cancer stage, (B) age range, (C) tumour size, (D) histological subtype (Lauren classification), (E) gender and (F) ethnicity. miR, microRNA.

negatives since this blood test served as a prescreening test for GC. Patients with positive test results will undergo endoscopy to confirm the diagnosis. Thus, this blood assay may reduce the reliance on endoscopy. Any test will have false negatives and it does not supplant clinical review and consideration for endoscopy if symptoms persist. It is not the intent of the 12-miR assay to replace endoscopic evaluation, we believe this assay provides an option for patients who might not be keen on initial endoscopic screening, and adds to the current cancer evaluation tool armamentarium, just as the stool DNA test is an option for colon cancer screening. Second, most controls in this study were patients with gastrointestinal symptoms referred to hospital clinics. Care should be taken when applying these findings to the general population. This 12-miR assay has obtained regulatory approval in Singapore and post-market surveillance data being gathered will clarify the PPV and NPV in different clinical settings, including in the general population. To date, the NPV in the general population is encouraging (data not shown). Furthermore, the study population is entirely Asian. Future studies in other populations should be considered. Finally, the roles that these GC-associated circulating miRNAs play in GC development and progression have not been defined through functional studies. Some of these miRNAs were shown to promote cancer metastasis and modulate tumour immune environment, additional mechanistic studies in cell and animal models are required.[29 30]

In conclusion, we have developed and validated a serum miRNA biomarker panel assay as a risk assessment tool for detecting GC. This assay is a useful adjunct in the armamentarium for cancer screening and has the potential to be a cost-effective mass screening tool for GC.

Table 4 Results of base-case analysis for mass screening for Singapore Chinese Males (50–75 years)

Cohort size[31]	482 469	
Total no of gastric cancer patients in the cohort	7241	
Compliance	45%	
	Screened	**Not screened**
No of subjects	217 111	265 358
Stage of diagnosis Stage1:2:3:4	26%: 17%: 24%: 32%	18%: 12%: 27%: 43%
Among those compliant to mass screening		
No of miRNA tests done to diagnose one gastric cancer patient	489	
No of endoscopes done to diagnose one gastric cancer patient	227	
Cost of saving 1 QALY* (US$)	44 531	

*Relative to the current practice of no-screening.
miRNA, microRNA; QALY, quality-of-life year.

Table 5 Cost-effectiveness of mass screening using 12-miR assay in conjunction with endoscopy

Strategy		Current practice: no screening	Mass screening with 3-yearly follow-ups
Cost	(US$)	173	533
Δ Costs	(US$)		+360
Efficiency	(QALY)	10.5032	10.5113
Δ Efficiency	(QALY)		+0.008
ICER	(US$/QALY)		44 531

ICER, Incremental cost-effectiveness ratio; miR, microRNA; QALY, Quality-adjusted life years.

Author affiliations
[1]Department of Surgery, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[2]Division of Surgical Oncology, National University Cancer Institute, Singapore
[3]Department of Surgery, National University Hospital, Singapore
[4]Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore
[5]Department of Medicine, Yong Loo Lin School of Medicine, National University Singapore, Singapore
[6]Department of Gastroenterology and Hepatology, National University Hospital, Singapore
[7]Department of Research and Development, MiRXES Lab, Singapore
[8]Diagnostics Development Hub, Accelerate Technologies Pte Ltd, Agency for Science Technology and Research, Singapore
[9]Yonsei Cancer Center, Songdang Institute for Cancer Research, Yonsei University College of Medicine, Seoul, Republic of Korea
[10]Saw Swee Hock School of Public Health, National University of Singapore, Singapore
[11]Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[12]Department of Surgery, Tan Tock Seng Hospital, Singapore
[13]Department of Gastroenterology and Hepatology, Tan Tock Seng Hospital, Singapore
[14]Singapore Immunology Network, Agency for Science Technology and Research, Singapore
[15]Department of Haematology-Oncology, National University Cancer Institute, Singapore
[16]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[17]Singapore Gastric Cancer Consortium, Singapore

Acknowledgements We would like to thank all our participants and the investigators from Yonsei Cancer Centre, NUH and TTSH hospitals, investigators from the Singapore Gastric Cancer Consortium (SGCC), the NUHS Tissue Repository, NUH Lab Medicine as well as the staff of all participating clinics, laboratories and research institutes for their contributions to the study.

Contributors JBYS, KGY, H-PT, JY, CTY, PCKG, JL, KPL, MH, CK, FZ and WPY participated in the design and performance of the study, review of results, analysis

and discussion. JBYS, CK, JL, SYR, HCC, JR, CKC, ST and AS enrolled subjects for the study and contributed clinical data. HPT, LZ, RZ, RK and JY analysed the data and RK performed the cost-effective analysis. The manuscript was drafted by JBYS, KGY, H-PT, LZ, YCT, RK, JY and reviewed by all authors. All authors read and approved the final manuscript.

Funding The Singapore Gastric Cancer Consortium (SGCC) is a national translational research group comprising clinicians and scientists working in gastric cancer research from academic medical centres, universities, hospitals and research institutes across Singapore. It receives funding from the National Research Foundation Singapore under its Translational and Clinical Research (TCR) Flagship Programme and Open Fund-Large Collaborative Grant (OF-LCG), administered by the Singapore Ministry of Health's National Medical Research Council. This study was supported by the Bedside & Bench grant (NMRC/BnB/0014b/2014) and the Translational and Clinical Research grant administered by Singapore Ministry of Health's National Medical Research Council (NMRC/TCR/009-NUHS/2013, NMRC/TCR/001-NUS/2007) and RIE2020 Centre Grant (CG) Programme (NMRC/CG/M005/2017_NCIS), as well as the COT and GAP grants administered by Singapore A*STAR Exploit Technology.

Competing interests KGY, JBYS, WPY, HPT, LZ, RZ and FZ were coinventors in the patent application 'Serum MicroRNA Biomarker for the Diagnosis of Gastric Cancer'. HPT, LZ and RZ are founders and shareholders of MiRXES. LZ, RZ and YCT are employees of MiRXES. HCC received grants from Lilly, GSK, MSD. Merck-Serono, BMS-Ono, Taiho outside the submitted work. The rest of authors declare no competing interests.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval The studies were approved by Domain Specific Review Board (DSRB) of the National Healthcare Group, Centralised Institutional Review Board (CIRB) of Singapore Health Services, Institutional Review Board (IRB) of National University of Singapore, Yonsei Cancer Center and Songdang Institute for Cancer Research.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

ORCID iDs
Feng Zhu http://orcid.org/0000-0001-6030-9251
Calvin Koh http://orcid.org/0000-0002-1756-3737
Jonathan Lee http://orcid.org/0000-0001-7065-8041
Kong-Peng Lam http://orcid.org/0000-0003-1558-1867
Khay-Guan Yeoh http://orcid.org/0000-0002-7802-4606

REFERENCES

1 Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016;66:7–30.
2 Van Cutsem E, Sagaert X, Topal B, *et al*. Gastric cancer. *Lancet* 2016;388:2654–64.
3 Choi KS, Jun JK, Suh M, *et al*. Effect of endoscopy screening on stage at gastric cancer diagnosis: results of the National cancer screening programme in Korea. *Br J Cancer* 2015;112:608–12.
4 Leung WK, Wu M-shiang, Kakugawa Y, *et al*. Screening for gastric cancer in Asia: current evidence and practice. *Lancet Oncol* 2008;9:279–87.
5 Shimada H, Noie T, Ohashi M, *et al*. Clinical significance of serum tumor markers for gastric cancer: a systematic review of literature by the task force of the Japanese gastric cancer association. *Gastric Cancer* 2014;17:26–33.
6 Ikeda F, Shikata K, Hata J, *et al*. Combination of Helicobacter pylori antibody and serum pepsinogen as a good predictive tool of gastric cancer incidence: 20-year prospective data from the Hisayama study. *J Epidemiol* 2016;26:629–36.
7 Huang Y-kai, Yu J-chun, Kang W-ming, *et al*. Significance of serum pepsinogens as a biomarker for gastric cancer and atrophic gastritis screening: a systematic review and meta-analysis. *PLoS One* 2015;10:e0142080.
8 Bartel DP. Metazoan microRNAs. *Cell* 2018;173:20–51.
9 Liang H, Gong F, Zhang S, *et al*. The origin, function, and diagnostic potential of extracellular microRNAs in human body fluids. *Wiley Interdiscip Rev RNA* 2014;5:285–300.
10 Cortez MA, Bueso-Ramos C, Ferdin J, *et al*. MicroRNAs in body fluids--the mix of hormones and biomarkers. *Nat Rev Clin Oncol* 2011;8:467–77.
11 Schwarzenbach H, Nishida N, Calin GA, *et al*. Clinical relevance of circulating cell-free microRNAs in cancer. *Nat Rev Clin Oncol* 2014;11:145–56.
12 Shigeyasu K, Toden S, Zumwalt TJ, *et al*. Emerging role of microRNAs as liquid biopsy biomarkers in gastrointestinal cancers. *Clin Cancer Res* 2017;23:2391–9.
13 Wei H, Pu K, Liu X-G, *et al*. The diagnostic value of circulating microRNAs as a biomarker for gastric cancer: a meta-analysis. *Oncol Rep* 2019;41:87–102.
14 De Guire V, Robitaille R, Tétreault N, *et al*. Circulating miRNAs as sensitive and specific biomarkers for the diagnosis and monitoring of human diseases: promises and challenges. *Clin Biochem* 2013;46:846–60.
15 Wong LL, Zou R, Zhou L, *et al*. Combining Circulating MicroRNA and NT-proBNP to Detect and Categorize Heart Failure Subtypes. *J Am Coll Cardiol* 2019;73:1300–13.
16 Huang KK, Ramnarayanan K, Zhu F, *et al*. Genomic and epigenomic profiling of high-risk intestinal metaplasia reveals molecular determinants of progression to gastric cancer. *Cancer Cell* 2018;33:137–50.
17 Miki K. Gastric cancer screening by combined assay for serum anti-Helicobacter pylori IgG antibody and serum pepsinogen levels - "ABC method". *Proc Jpn Acad Ser B Phys Biol Sci* 2011;87:405–14.
18 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stats Soc Ser B* 1995;57:289–300.
19 Xiong M, Fang X, Zhao J. Biomarker identification by feature wrappers. *Genome Res* 2001;11:1878–87.
20 Cohen JF, Korevaar DA, Altman DG, *et al*. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016;6:e012799.
21 Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: a comparison of resampling methods. *Bioinformatics* 2005;21:3301–7.
22 Kawaguchi T, Komatsu S, Ichikawa D, *et al*. Circulating microRNAs: a next-generation clinical biomarker for digestive system cancers. *Int J Mol Sci* 2016;17. doi:10.3390/ijms17091459. [Epub ahead of print: 01 Sep 2016].
23 Kinoshita T, Yip KW, Spence T, *et al*. MicroRNAs in extracellular vesicles: potential cancer biomarkers. *J Hum Genet* 2017;62:67–74.
24 Huang Y-K, Yu J-C. Circulating microRNAs and long non-coding RNAs in gastric cancer diagnosis: an update and review. *World J Gastroenterol* 2015;21:9863–86.
25 Link A, Kupcinskas J. Micrornas as non-invasive diagnostic biomarkers for gastric cancer: current insights and future perspectives. *World J Gastroenterol* 2018;24:3313–29.
26 Sawaki K, Kanda M, Kodera Y. Review of recent efforts to discover biomarkers for early detection, monitoring, prognosis, and prediction of treatment responses of patients with gastric cancer. *Expert Rev Gastroenterol Hepatol* 2018;12:657–70.
27 Cissell KA, Deo SK. Trends in microRNA detection. *Anal Bioanal Chem* 2009;394:1109–16.
28 Loeb S, Vonesh EF, Metter EJ, *et al*. What is the true number needed to screen and treat to save a life with prostate-specific antigen testing? *J Clin Oncol* 2011;29:464–7.
29 Le MTN, Hamar P, Guo C, *et al*. miR-200-containing extracellular vesicles promote breast cancer cell metastasis. *J Clin Invest* 2014;124:5109–28.
30 Eichmüller SB, Osen W, Mandelboim O, *et al*. Immune modulatory microRNAs involved in tumor attack and tumor immune escape. *J Natl Cancer Inst* 2017;109. doi:10.1093/jnci/djx034. [Epub ahead of print: 01 Oct 2017].
31 Department of Statistics, Ministry of Trade and Industry. *Population trends, ISSN 1793-2424*. Republic of Singapore, 2016. http://www.singstat.gov.sg

Supplementary Material

Supplementary Methods

Analytical Validation of miRNA Assays

The analytical performance of the miRNA assays used for gastric cancer biomarker identification was evaluated. We first evaluated the analytical specificity of these assays by conducting a cross-reactivity test of miRNA assays against 9 highly homologous let-7 family members (**Figure S2A**), a design routinely used to evaluate assay specificity. The let-7 family assays were able to discriminate homologous sequences with even single nucleotide differences, e.g. let-7a assay showed 100% detection of let-7a target and only 0.8% cross-reactivity against let-7c target. Secondly, we evaluated the reproducibility of the miRNA assays by measuring 200 circulating miRNAs in 30 control and cancer serum specimen in two independent labs (**Figure S2B**). After normalization of technical and biological variations using the multi-layered controls illustrated in **Figure S1**, these assays demonstrated encouraging concordance of 0.95-0.98 in all 30 clinical samples. Lastly, we evaluated the analytical sensitivity of the miRNA assays (**Figure S2C**). Constrained by the small size, miRNA assay performances can be highly variable across different miRNA targets, especially miRNAs with higher AT content. We selected 8 commonly studied miRNAs with low to high AT content (36.4% - 63.6%) and compared the analytical sensitivity and dynamic range of the MiRXES miRNA assays against the well known Taqman probe based assays. The miRNA assays used for gastric cancer biomarker discovery demonstrated consistent amplification and detection of all 8 miRNA targets across at least 7 logs of dynamic range where the probe-based assays showed less consistent performance, especially against miRNA targets with higher AT content. Overall, these validation studies demonstrated good analytical performance of the assays and the workflow, and warrant their use for biomarker discovery.

Laboratory Procedures for miRNA Expression Quantification in Discovery and Verification Phases

Spike-In Controls for RT-qPCR Workflow

To monitor and normalize technical variations in RNA isolation efficiency, a set of 3 proprietary synthetic miRNAs were spiked into the sample lysis buffer (Qiazol) at high, medium and low concentrations. To monitor and normalize technical variations in subsequent RT and qPCR reactions, a second set of 3 proprietary synthetic miRNAs were then spiked into each isolated sample RNA at high, medium and low concentrations. A 6-log serial dilution of synthetic templates (10^7 to 10^2 copies) for each miRNA, non-template control (nuclease-free water spiked with MS2) and reference human serum RNA were concurrently reversed-transcribed and quantified by qPCR with each isolated serum RNA sample. These control measures facilitated monitoring and normalization of technical variations in pipetting and assay efficiency in RT, cDNA amplification, and qPCR.

Determination and Normalization of Absolute miRNA Expressions

Upon completion of RT-qPCR, Ct values were determined using the ViiA 7 RUO software (Thermo Fisher Scientific Inc, USA) with automatic baseline setting and a threshold of 0.5. Absolute expression of each miRNA in patient serum was determined through intra-polation of synthetic miRNA standard curves and corrected for RT-qPCR efficiency variation using spike-in RNAs. The miRNA expression of each sample was further normalized using 6 endogenous reference miRNAs independently identified using the geNorm and NormFinder reference gene algorithms [1, 2]. The miRNA expression profiles normalized using the 6 reference miRNAs were similar to that normalized by global mean expression of all miRNA quantified. Absolute expression of miRNAs were log2 transformed for subsequent statistical analysis and optimization of multivariate biomarker panel.

Multivariate Analysis For Constructing Multi-miR Panels

A linear support vector machine (SVM) was used to construct the multi-variant biomarker panels and the associated algorithm that classified cancer and control groups with highest AUC. Multiple iterations of four-

So JBY, *et al. Gut* 2020;0:1–9. doi: 10.1136/gutjnl-2020-322065

fold cross-validation (matched by sex, cancer subtype and disease stage) were conducted to evaluate the performance of these panels. All calculations were performed using Matlab® software (MathWorks, USA).

Laboratory Procedures for 12-miR Multi-Target Assay in Validation Phase

The assay involved 3 steps: (1) RNA isolation from serum samples; (2) cDNA synthesis; and (3) Detection of miRNAs by quantitative PCR (qPCR). Extraction of RNA was performed by combining phenol/guanidine-based lysis of serum sample and silica-membrane-based purification of total RNA. During cDNA synthesis, 12 miRNA targets from each specimen were converted into cDNAs using 12 corresponding miRNA-specific stem-loop-based reverse transcription primers in a single reaction on a Veriti Dx thermocycler (Thermo Fisher Scientific Inc, USA). At the qPCR step, each miRNA target was amplified by a sequence-specific forward PCR primer and a hemi-nested sequence specific reverse PCR primer and detected using SYBR Green I dye in single-plex reactions on a Quantstudio Dx (384-well) real-time qPCR instrument (Thermo Fisher Scientific Inc, USA). Ct values of the 12 biomarker and reference miRNAs were exported using the QuantStudio Dx Software v1.0.1 (Thermo Fisher Scientific Inc, USA) and converted into a single numerical score using a validated, prespecified logistic-regression algorithm through the GASTROSmart Software (MIRXES Pte Ltd, Singapore). In each assay run, 13 patient specimens were processed concurrently with 2 quantitative reference specimen and 1 negative control specimen, which served as quality control and inter-run normalizers.

Cost-effectiveness analysis

We examined the cost-effectiveness of implementing the miRNA biomarker panel as a screen before endoscopy in a proposed national screening program in Singapore. Our study focusses on the cohort of Singaporean Chinese males, age 50-75 years, who are at an intermediate risk of gastric cancer, and compare the proposed mass screening program with the current pattern of gastric cancer diagnosis without screening. Chinese population carry ~90% of gastric cancer disease burden in Singapore with males at a 30% higher risk of gastric cancer than females [3, 4]. With the cancer incidence rising sharply after the age of 50 years[4], this subgroup with intermediate gastric cancer risk has a 4 times higher annual incidence rate than the general population. We estimate the quality-adjusted life years (QALY), costs per person, incremental cost-effectiveness ratio (ICER) and also the benefits of early cancer diagnosis and reduced mortality achieved by implementing the mass screening program.

Detailed research methodology of cost-effectiveness analysis is as follows:

- **Target Population:** The analysis is performed on the cohort of Singaporean Chinese males aged 50-75 years.
- **Interventions Compared**
  The two interventions compared are:
  1. Current practice of no screening.
  2. Mass screening program using miRNA -test, followed with test-positive patients undergoing a confirmatory upper-endoscopy and biopsy and test-negative subjects to be followed up 3-yearly.

- **Methodology**
  Markov decision model was built in Microsoft Excel 2010 to compare the two interventions in the target population by analyzing in a closed cohort setting (**Figure S7**). Model was populated using local and published data with the cohort size estimated from 2016 population census [5]. With a healthcare system perspective, a 25 years' time horizon was analyzed with subjects exiting the model at the age of 75 years. Subjects were expected to be in one of the five health conditions – healthy (cancer-free), TNM Stage 1, TNM Stage 2, TNM Stage 3 and untreatable terminal stage (Stage 4). Early or advanced stage patients (stage 1, 2, and 3) received curative treatment with a stage specific cancer recurrence possibility after a mean duration of 2 years [6, 7], while terminal

cancer patients (stage 4) received only palliative care with a conditional life expectancy of 1 year [6]. As prognosis of the cancer recurrence is poor, patients diagnosed with recurrence were assumed untreatable (equivalent to stage 4) and were given palliative care. Only gastric cancer related mortality was compared as the background mortality due to natural or other causes was expected to be similar in both the scenarios.

The current practice of no screening evaluates the costs, health impacts and mortality as per the current diagnosis rate of gastric cancer in this specific population cohort. We used the published age-specific annual incidence rates of gastric cancer and stage of diagnosis among Chinese Males in Singapore. In the current practice of no-screening we did not account for the cost of false positive endoscopies and the diagnostic expenses of only the true cancer cases was considered, which was a conservative assumption favoring no-screening, similar to the assumption in earlier studies [8, 9]. The proposed mass screening program on the other hand was expected to screen the compliant cohort, identify the cancer cases early due to regular screening and computes the cost, health impacts and mortality accordingly. The subjects tested negative in the screening program will include both healthy cases and missed cancer cases. The missed cancer patients were expected to experience the consequences of treatment delays – disease progression, impact on cost and quality-of-life and an increased mortality, as the cancer would progress in them undiagnosed and untreated and the healthy cases are expected to remain healthy with a possibility of developing gastric cancer in future. Cancers missed in the mass screening program were considered to progress to advanced stages and are expected to be diagnosed at stage 4 due to presentation of symptoms in clinics where they are investigated by endoscopy and biopsy. A 1-year progression time was estimated between the consecutive cancer stages, i.e. a missed stage 1 cancer is expected to progress to stage 2 and then to stage 3 and stage 4 with a one year gap each between the successive stages. Stage 4 patients which were missed in diagnosis were expected to be diagnosed after a mean time of 2 months due to worsening of symptoms.
The compliance rate for mass screening was assumed to be 45% as per the reported compliance in national gastric cancer screening programs in Korea [10] with the non-compliant group expected to behave similar to the current strategy of no-screening. As the miRNA test is simpler to administer and potentially cheaper than the currently used screening methods of UGIS, X-ray or endoscopy, it is hoped to improve the population compliance rate. Thus the performance of the mass screening program across a range of compliance rates (45% - 100%) was also evaluated. All costs quoted in US dollars have been calculated based on the exchange rate of $1.38 Singapore dollars to 1 USD as per exchange rates in July 2017. All costs and health benefits were discounted at an annual rate of 3%.

- **Scenario and sensitivity analysis:** The cost advantages and non-invasive nature of miRNA testing may increase patient compliance with screening relative to current technologies. Scenarios that capture a range of improved compliance rates (45%–100%) were modeled to evaluate the possible impact on early diagnosis (**Figure S5**). An extensive sensitivity analyses was conducted by varying the values of key parameters—endoscopy cost, miRNA test costs, miRNA test specificity/sensitivity by cancer stages (stages 1, 2, 3, 4), QoL values by cancer stages (stage 1, 2, 3, 4), cancer recurrence rates by the stage at diagnosis and average annual incidence of gastric cancer—to evaluate model robustness at a Willingness-to-Pay threshold of 50,000 USD/ QALY (**Figure S8, 9, 10**).

- **Treatment Protocols for Cancer Treatment and Related Costs:**
Stage-specific treatment protocols and average medical expenditures for gastric cancer were obtained from the National University Hospital and expert opinions of clinicians based on current

practices in Singapore (Supplementary Table11). Patients diagnosed with gastric cancer undergo staging investigation, which includes Computerized Tomography (CT), Chest X-Ray (CX-R), Endoscopic Ultrasound (EUR) and a specialist consultation (including the cost of nurse counseling and an estimated round-trip transport). Curative treatment administered to stage 1, 2, and 3 cancer patients includes surgery (total/ partial gastrectomy) and hospital stay of 12 days. Stage 3 patients undergo an additional chemo-radiotherapy (5 follow-ups) and radiotherapy sessions (5 sessions/ week for 5 weeks). Palliative care for stage 4 patients includes bypass surgery (30%), endoscopic stenting (6%), palliative chemotherapy-5 sessions (16%) and conservative treatment (2x specialist visits) (48%) with an appropriate hospital stay (12 days in surgery cases and 2.5 days for cases with no surgery). Patients are also expected to adhere to follow up visits (average 2.2 visits/ 5 years) and repeat CT and CXR (average 1.5per year for 5years after the diagnosis).
The miRNA panel test cost in Singapore has been assumed to be USD 30 with an additional 10% for handling and administrative purposes. However, the cost of organizing mass screening has not been included. Costs and QALYs were presented on a present-value basis, with an annual discount rate of 3%. All the diagnosed cases are expected to undergo a biopsy examination. Total costs have been evaluated inclusive of Goods and Services Tax (GST) and without considering any government subsidy.

- **Quality of life values:**
  Stage-specific EQ-5D quality-of-life (QoL) index measures were obtained from a previous local study [11] performed on Chinese gastric cancer patients in National University Hospital, Singapore (**Table S6**). A diagnosed patient is expected to be immediately started on treatment, and experience the diagnosed stage-specific QoL for 1 year with a 6-month additional decrease in QoL due to the initial surgery referred as disutility. After one year of treatment, the patient is expected to enjoy a QoL equivalent to an asymptomatic patient (similar to stage 1 cancer) for the remaining time until faced with any recurrence, which would subsequently drop the QoL to stage 4 equivalent (**Table S6**).

- **Test Characteristics:**
  Test characteristics for diagnostic endoscopy with biopsy for the suspected cases (sensitivity: 93%, specificity: 100%) has been resourced from a study evaluating diagnostic accuracy through a retrospective study among gastric cancer patients [12]. Biopsy is believed to be perfect with 100% sensitivity and specificity. The miRNA stage –specific sensitivity and specificity as estimated from the Singapore Discovery Cohort have been considered as the base-case value.

- **Estimation of population prevalence of undiagnosed gastric cancer:**
  As the study aim to identify the benefits of early diagnosis of gastric cancer, it is essential to calculate the population prevalence of undiagnosed gastric cancer cases in the target group. The current annual age-specific incidence rate is 57 cancers per 100,000 in this population cohort [4] with a stage specific distribution of stage 1: 2:3: 4:: 18%:12%:27%:43% . Based on the assumption of 1 year time for progression of cancer from one stage to another, undiagnosed cancer prevalence in the population cohort (stage 1 and higher) was evaluated individually before every mass screening follow-up. Also, the stage 1 and 2 cancers which currently develop and are diagnosed in between the follow-up years are expected to continue to be diagnosed in both the strategies.

So JBY, *et al. Gut* 2020;0:1–9. doi: 10.1136/gutjnl-2020-322065

Supplementary Figures

Figure S1. Multi-layered control measures for absolute quantification of miRNA expression

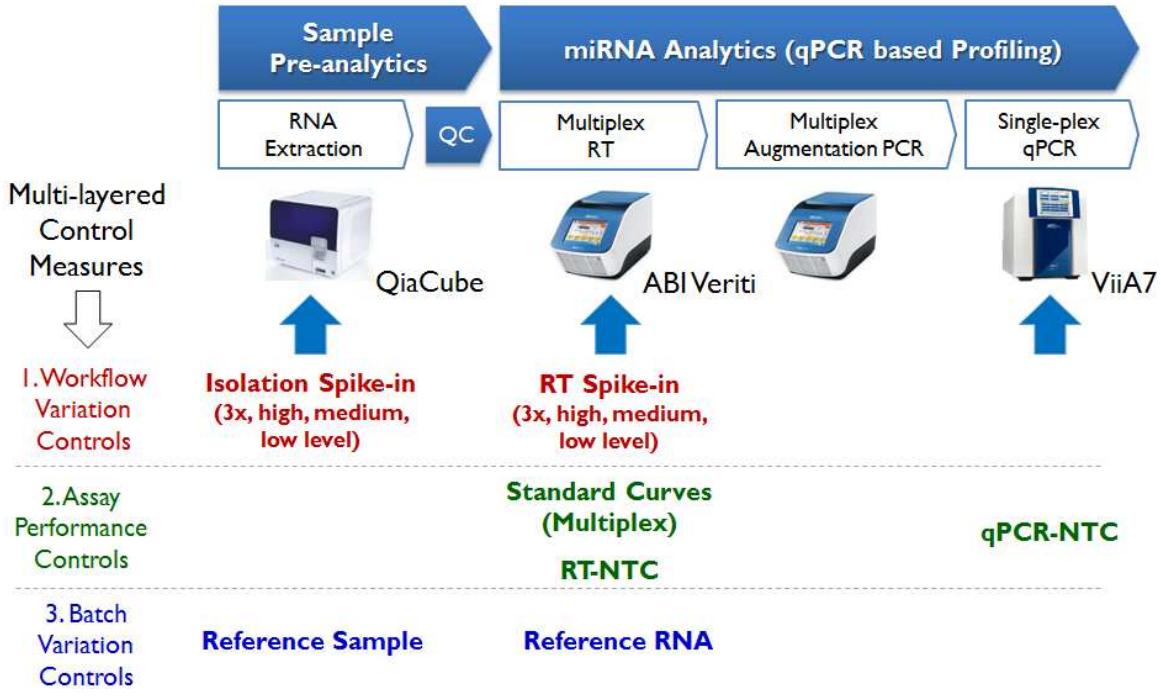


Figure S2. Analytical Validation of MIRXES miRNA RT-qPCR Assays. (A) Analytical Specificity Test. Table below shows the relative detection when assays are challenged with mismatched let-7 family members. **(B) Analytical Reproducibility Test**. Correlation of the expression profiles of 200 miRNAs quantified in 15 control and 15 gastric cancer sera in two independent laboratories (R2> 0.95) using MIRXES miRNA RT-qPCR assays. **(C) Analytical Sensitivity Test.** Comparison with Probe Based Assays. Consistency in the analytical sensitivity of the MIRXES assays was demonstrated by amplification and detection of a 7-log serial dilution of the synthetic templates of 8 miRNAs with low to high AT content. In contrast, probe based assays showed poorer consistency across miRNAs with different AT content.

Figure S2A

	Target	Assay relative detection								
		let-7a	let-7b	let-7c	let-7d	let-7e	let-7f	let-7g	let-7i	miR-98
UGAGGUAGUAGGUUGUAUAGUU	let-7a	100.00%	0.00%	0.10%	1.20%	0.00%	0.50%	0.00%	0.00%	0.00%
UGAGGUAGUAGGUUGUGUGGUU	let-7b	0.00%	100.00%	2.20%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
UGAGGUAGUAGGUUGUAUGGUU	let-7c	0.80%	1.60%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
AGAGGUAGUAGGUUGCAUAGUU	let-7d	0.40%	0.00%	0.10%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
UGAGGUAGGAGGUUGUAUAGUU	let-7e	0.10%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
UGAGGUAGUAGAUUGUAUAGUU	let-7f	0.10%	0.00%	0.00%	0.00%	0.10%	100.00%	0.00%	0.00%	0.00%
UGAGGUAGUAGUUUGUACAGUU	let-7g	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
UGAGGUAGUAGUUUGUGCUGUU	let-7i	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
UGAGGUAGUAAGUUGUAUUGUU	miR-98	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%

Figure S2B

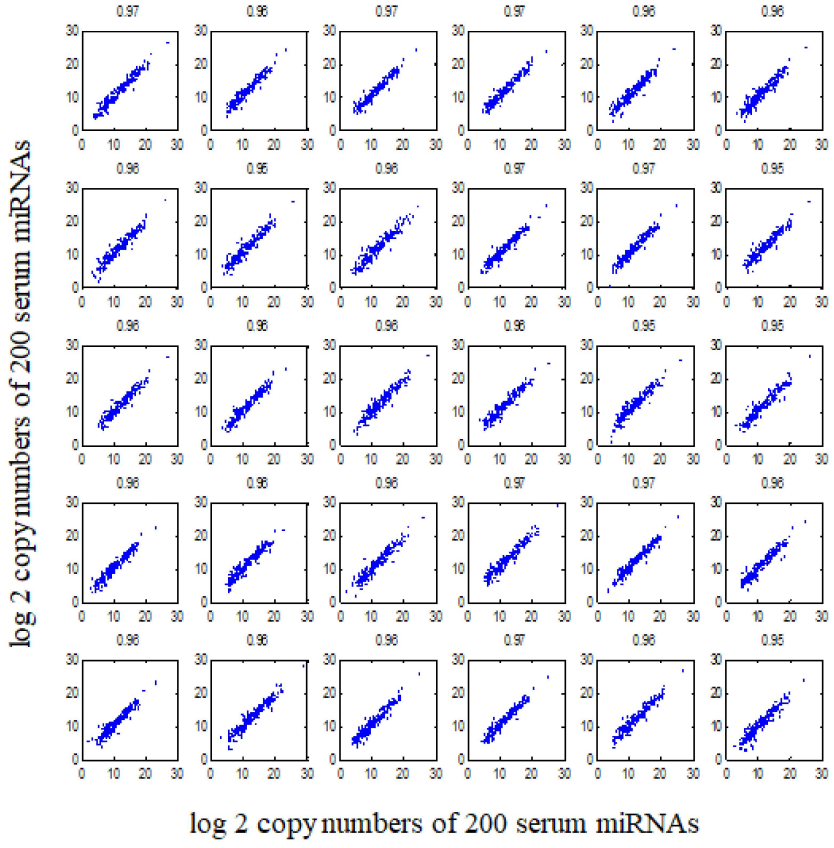
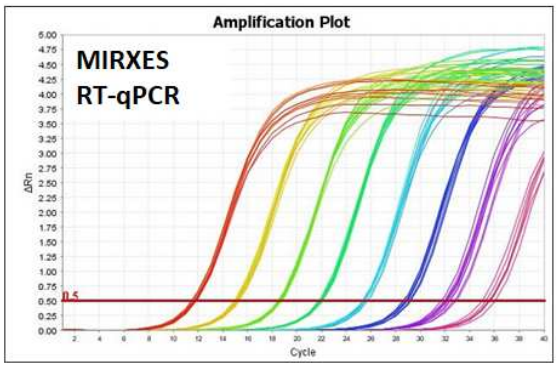


Figure S2C

#	miRNA	Sequence	AT%
1	hsa-miR-99b-5p	CACCCGUAGAACCGACCUUGCG	36.4%
2	hsa-miR-133a-3p	UUUGGUCCCCUUCAACCAGCUG	45.5%
3	hsa-miR-593-3p	UGUCUCUGCUGGGGUUUCU	47.4%
4	hsa-miR-151a-5p	UCGAGGAGCUCACAGUCUAGU	47.6%
5	hsa-miR-20b-5p	CAAAGUGCUCAUAGUGCAGGUAG	52.2%
6	hsa-miR-122-5p	UGGAGUGUGACAAUGGUGUUUG	54.5%
7	hsa-miR-377-3p	AUCACACAAAGGCAACUUUUGU	59.1%
8	hsa-miR-451a	AAACCGUUACCAUUACUGAGUU	63.6%



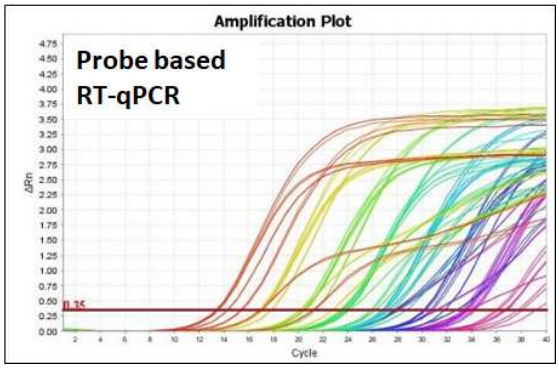


Figure S3. (A) Top up-regulated and down-regulated miRNAs in gastric cancer. The boxplots and ROC curves of top-ranked (based on AUC) up-regulated (miR-142-5p) and down-regulated (miR-99b-5p) miRNAs in all the gastric cancer subjects (regardless of subtypes and stages) compared to normal subjects; the expression levels (copy/ml) were presented in log2 scale. The boxplot presented the 25th, 50th, and 75th percentiles in the distribution of expression levels. AUC: area under the ROC curve. **(B) Top up-regulated and down-regulated miRNAs between normal and various stages of gastric cancers.** The ROC curves of top-ranking (based on AUC) up-regulated (row 1) and down-regulated (row 2) miRNAs in the various stages of gastric cancers subjects compared to the normal subjects. AUC: area under the ROC curve.

Figure S3A

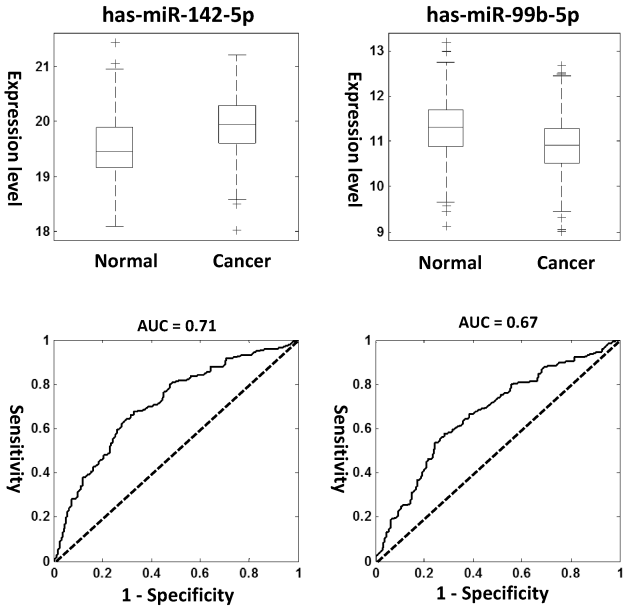


Figure S3B

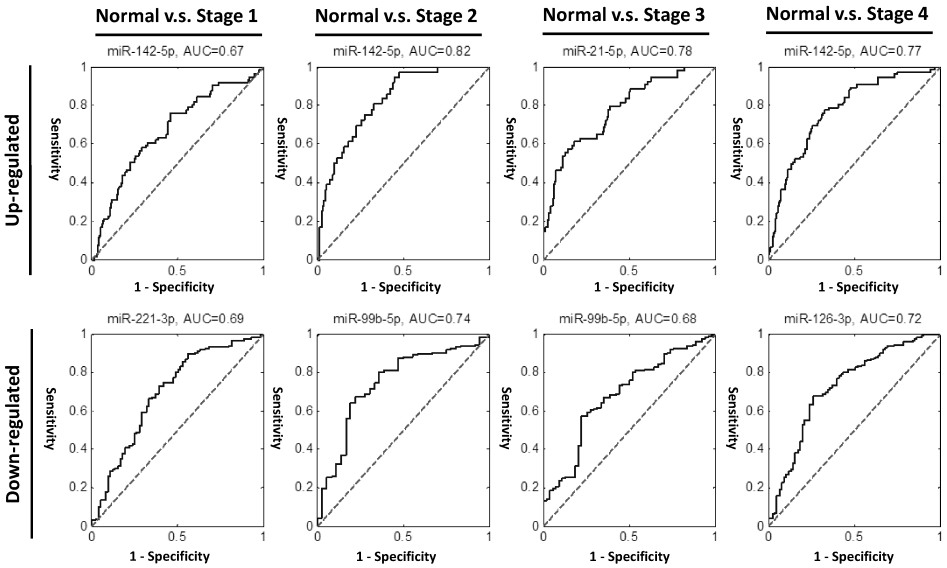

Normal v.s. Stage 1
Normal v.s. Stage 2
Normal v.s. Stage 3
Normal v.s. Stage 4
Up-regulated
Down-regulated
miR-142-5p, AUC=0.67
miR-142-5p, AUC=0.82
miR-21-5p, AUC=0.78
miR-142-5p, AUC=0.77
miR-221-3p, AUC=0.69
miR-99b-5p, AUC=0.74
miR-99b-5p, AUC=0.68
miR-126-3p, AUC=0.72
Sensitivity
1 - Specificity

Figure S4. Differentially Expressed miRNAs between Various Gastric Cancer Subtypes. Boxplot of 7 miRNAs with p-values lower than 0.05 for three different subtypes based on two-way Anova test after false discovery rate (FDR) correction (Bonferroni method). The p-values were shown above the miRNAs. The boxplot presented the 25th, 50th, and 75th percentiles in the distribution of log2 scale expression levels (copy/ml). For each miRNA, the significant levels between various subtypes were calculated based on Bonferroni adjustment to compensate for multiple comparisons. *: p-value < 0.05; **: p-value < 0.01; ***: p-value < 0.001. Five (hsa-miR-27a-3p, hsa-miR-338-5p, hsa-miR-181d, hsa-miR-146b-5p, hsa-miR-30e-3p) of these 7 miRNAs were found to be up-regulated and 2 were found to be down-regulated (hsa-miR-21-3p, hsa-miR-1226-3p) in the diffuse subtype when compared to the intestinal subtype. In addition, the expression level of these 7 miRNAs in the mixed subtype were found to be similar to the diffuse subtype except for hsa-miR-146b-5p, where the expression level in diffuse subtype was higher than the other two subtypes (middle lower panel).

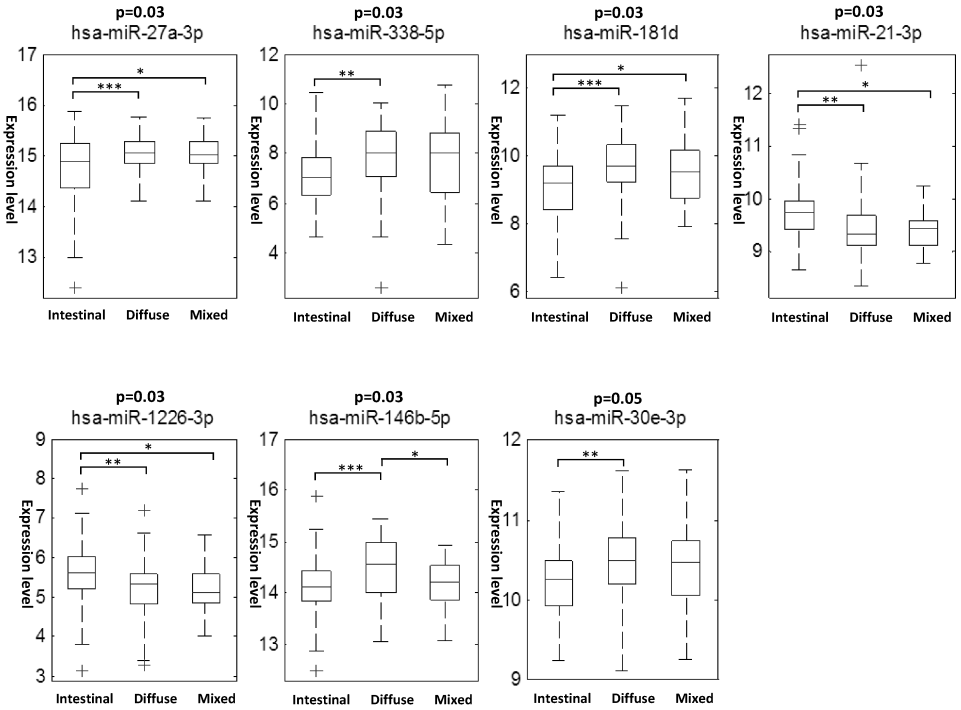


Figure S5. The impact of population compliance in screening program on early cancer diagnosis. With increasing compliance in the mass screening program, more cancer cases are expected to be diagnosed in early stages as compared to late stages. The figure below represents the expected pattern of early diagnosis with higher compliance rates. The ICER which is the cost spent to gain 1 additional QALY in the life of patient is independent of the compliance rate and is equivalent to USD 28,931/ QALY in all scenarios. *Abbreviations used- QALY: Quality-adjusted life years, ICER: Incremental cost-effectiveness ratio, Mn: Million, USD: United States Dollars*

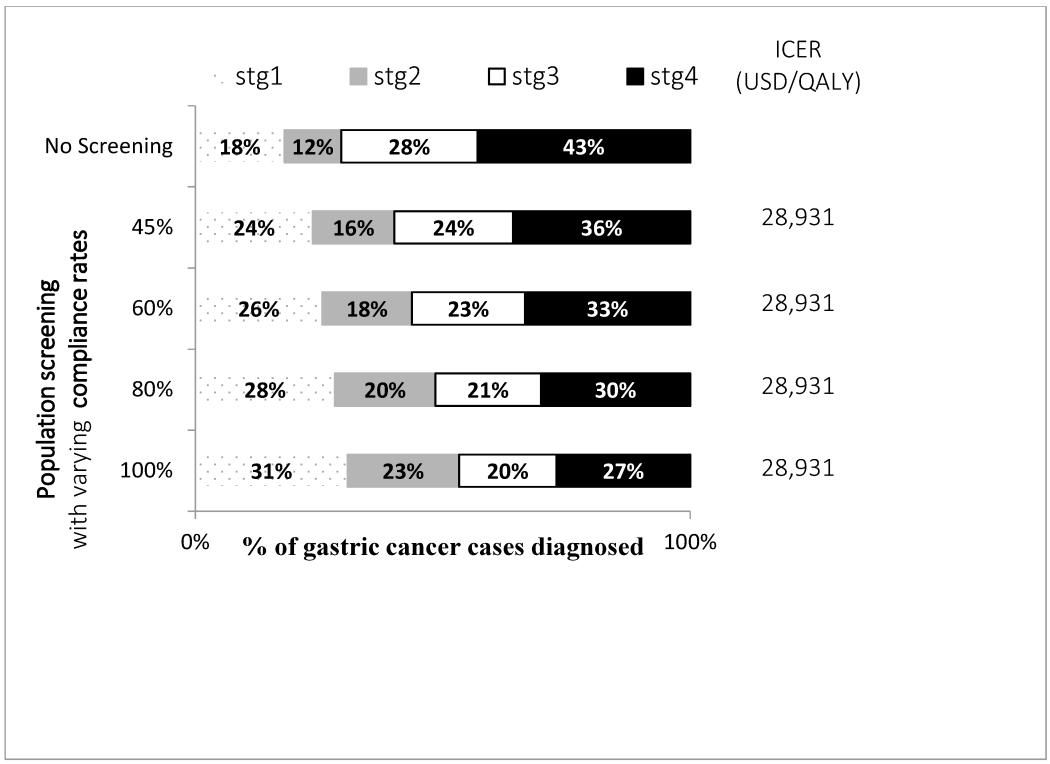


Figure S6. Comparison of Serum miRNA Biomarkers for Gastric, Breast and Ovarian Cancers. Three independent studies have been conducted to investigate serum miRNA expression changes between gastric, breast and ovarian cancers with their corresponding control populations. All measurements were performed using identical assays and workflows. The overlaps of up- and down-regulated serum miRNA biomarkers for these three cancers were presented in the venn diagram below. While there are some overlaps among the 3 cancers, distinct miRNAs changes specific to each cancer were observed. The heatmap below illustrates distinct serum miRNA expression in breast and gastric cancer patients.

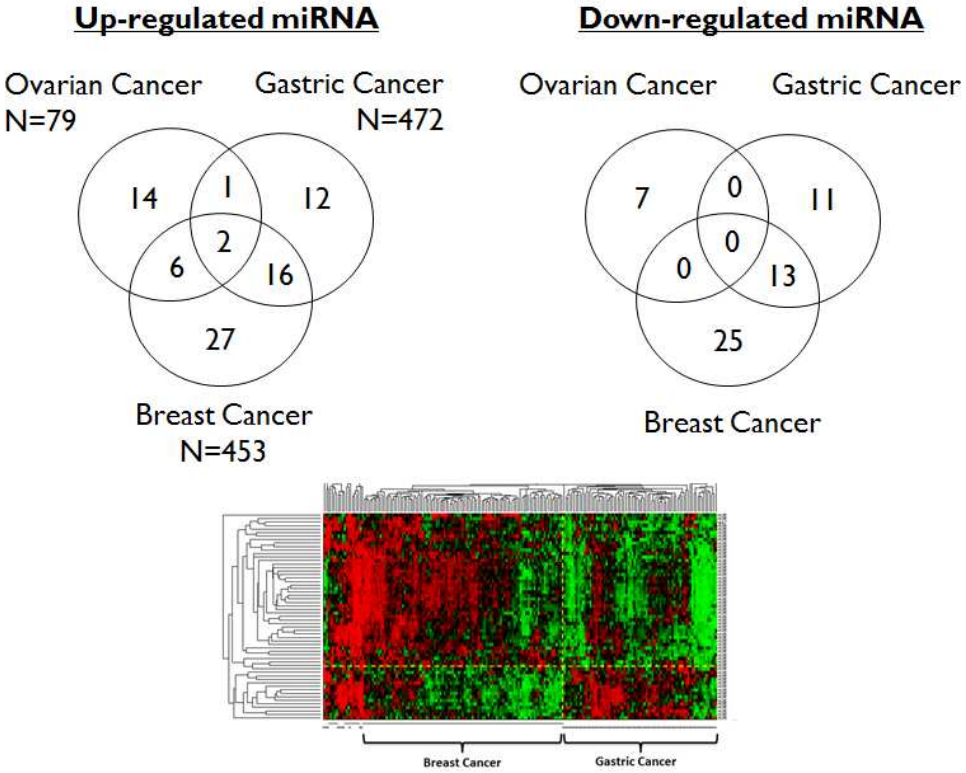


Figure S7. Markov-Decision Tree Model for evaluating the cost-effectiveness. The Markov decision model built has a 25-year time horizon and represents the movement of subjects from one health state to another in both the strategies. The patients who are diagnosed with cancer are treated as per the stage of diagnosis. Cancer recurrence is considered the only reason of treatment failure with all the recurrent cases being fatal. Subject is expected to exit the cohort at age 75. Health and cost parameters corresponding to each state are indicated at every step. The Markov model considers a 3% discount rate for both cost and health benefits and calculates values in their net present value.

Figure S7(A) Decision Tree Model for Strategy 1 – No screening. In current practice the subjects are diagnosed in clinic and treated as per the stage of diagnosis. The subjects in the population cohort would continue living healthy until being diagnosed as per the annual incidence rate reported in the Singapore cancer registry.

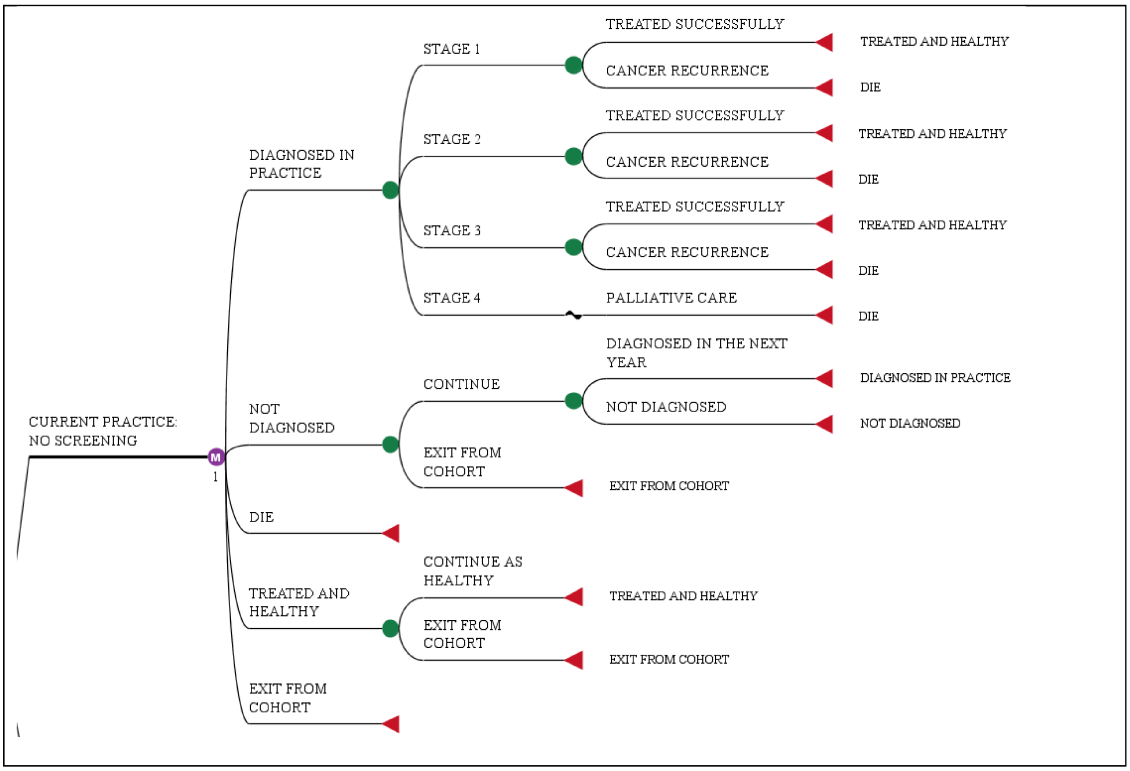


Figure S7(B) Decision Tree Model for Strategy2: mass screening with miRNA followed by endoscopy for diagnosis confirmation for test positive subjects and a 3-year follow up for test negative subjects. The subjects compliant to the mass screening program are made to undergo a miRNA-based blood test. If the miRNA test is positive, the subjects will undergo endoscopy and biopsy to confirm the cancer and if the miRNA test is negative, subjects are considered healthy until new cancers are diagnosed over time. Among this healthy group, there would be cancer cases missed by miRNA test. The missed cancer cases would progress to advanced stages and are expected to be diagnosed with symptoms at stage 4 in the clinic. The remaining healthy cases group is followed up with miRNA test every 3 years.

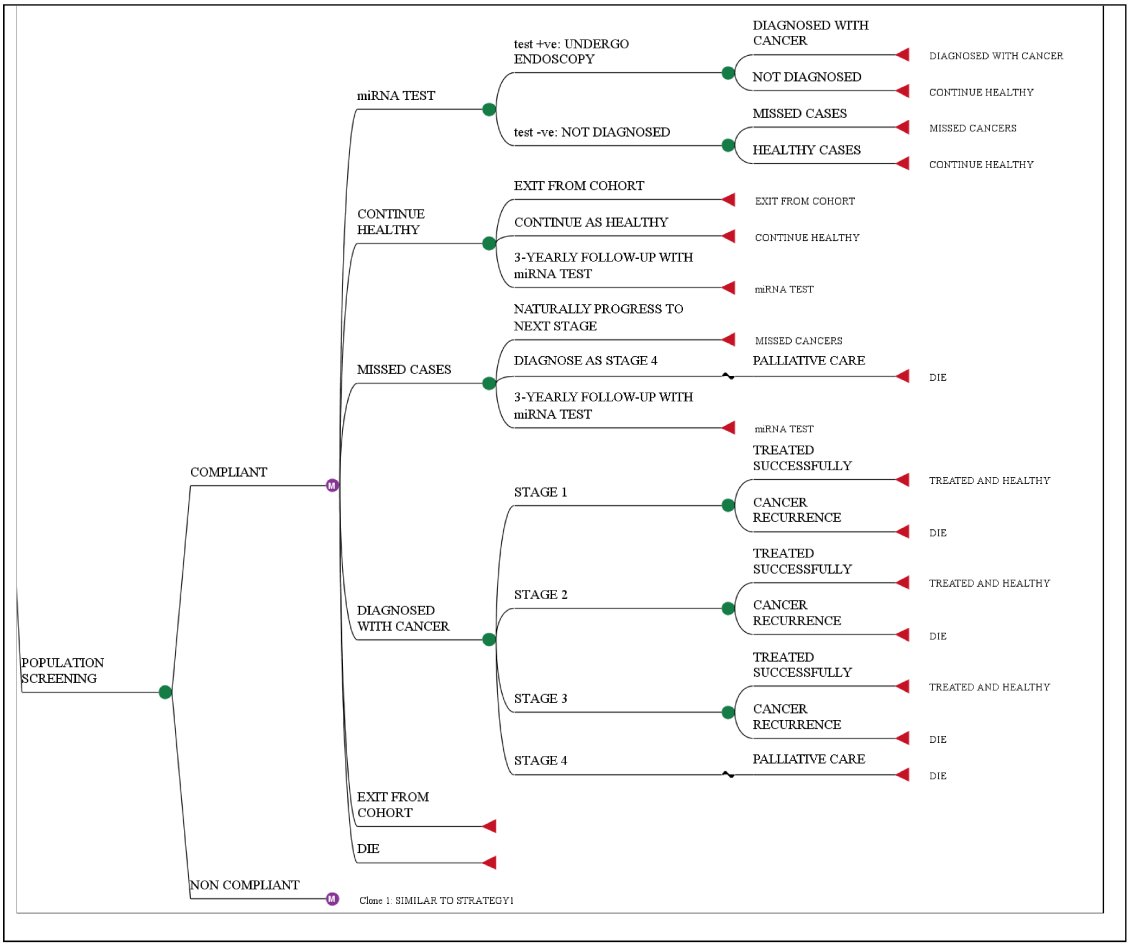


Figure S8. Sensitivity analysis for mass screening of Singaporean Chinese Males (50-75 years). We have performed one-way sensitivity analysis of many key variables to identify the impact of variable uncertainty on the Incremental Cost Effectiveness Ratio (ICER). The figure below shows all variables with their sensitivity ranges on Y axis and the ICER values on X axis. The range of values that were examined is shown in parentheses, with the value giving the lower ICER listed first. The graph represents the possible variation in ICER due to variable uncertainty, with the most significant variables at top. The solid vertical line indicates the ICER of 28,931 USD/QALY for the base-case scenario while the dash line indicates the threshold of ICER 50,000 USD/QALY. Three significant variables were identified which are: miRNA test cost, specificity of miRNA test and sensitivity of miRNA test for stage 1 patients. *Abbreviations used- QALY: Quality adjusted life years, ICER: Incremental cost-effectiveness ratio*

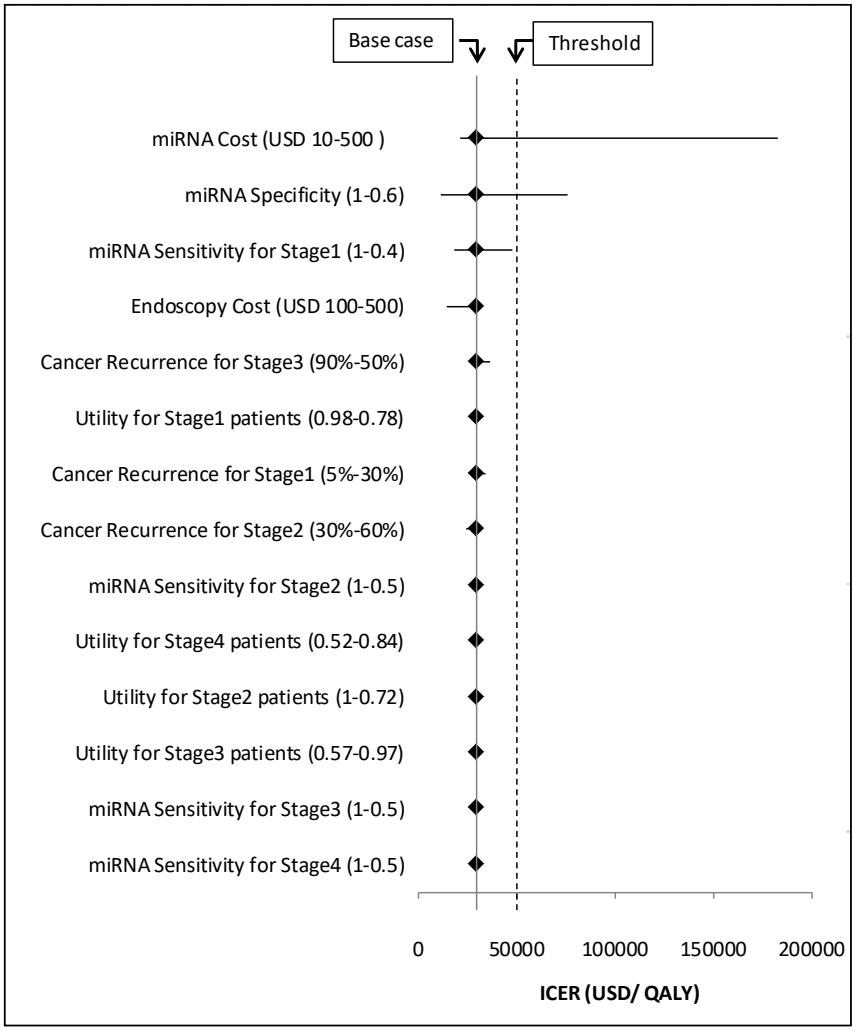


Figure S9. One-way sensitivity analysis for variables found significant for cost-effectiveness evaluation. The variation in ICER within the range of values of the three significant variables has been shown below individually. Also, their threshold value, i.e. value at which the ICER is equivalent to USD 50,000/QALY has been highlighted. The strategy would be cost-effective only with an ICER < 50,000 USD/ QALY. For each sensitivity analysis below, it is assumed that the rest of variables remain constant as described in base-case scenario.

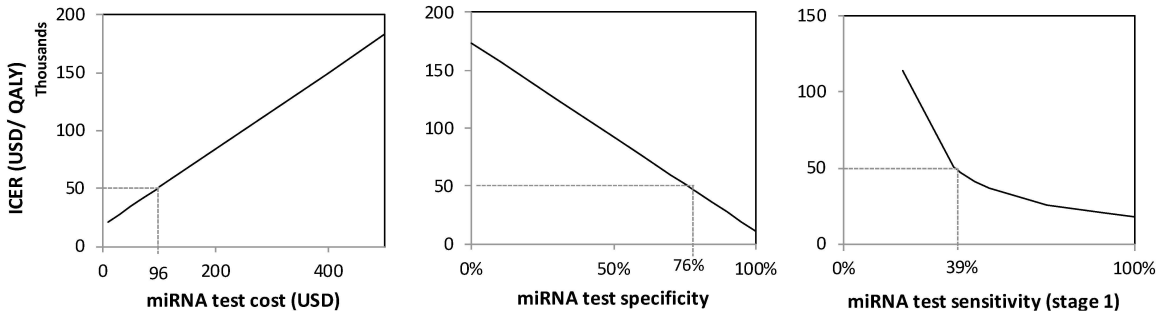


Figure S10. Sensitivity analysis of the cost-effectiveness of screening program with varying gastric cancer incidence. This sensitivity analysis has been performed to identify the cancer incidence which makes screening programs in the population cost-effective. The graph below reports the ICER at the different incidence rates ranging from 0.01% - 0.5%. The incidence rate reported is the average annual incidence for the target population. The analysis found the screening program to be cost-effective at an average incidence higher than 0.05% based on the cost-effectiveness threshold of US$50,000/QALY. The target population of Singaporean Chinese males (50-75 years) is estimated to have an average annual gastric cancer incidence of 0.057% at the beginning of the analysisas per the 2016 population cohort statistics. With ageing, the average incidence rate of the cohort has been considered to increase.

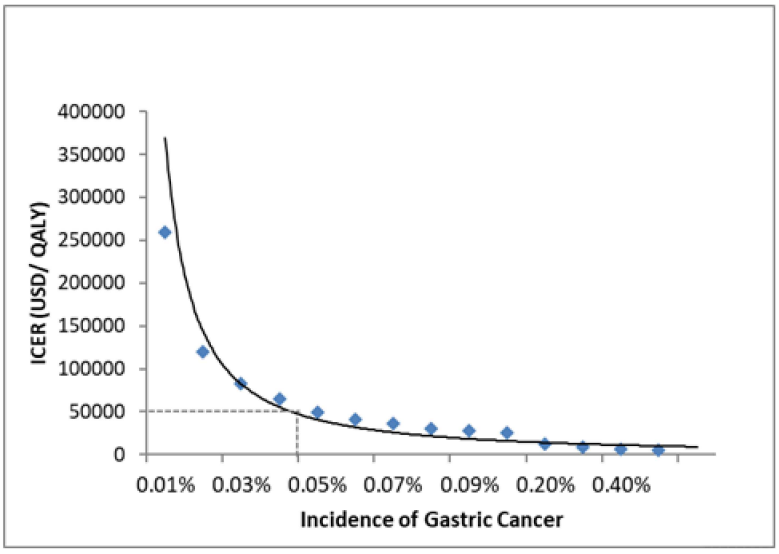


Table S1. Identity and Sequence of 191 Reliable Detected Mature miRNA. 191 mature miRNA were reliable detected in the serum samples. The definition of "reliably detected" was that at least 90% of the serum samples had a concentration higher than 500 copies per ml. The miRNAs were named according to the miRBase V18 release.

Identity	Sequence
hsa-miR-99b-5p	CACCCGUAGAACCGACCUUGCG
hsa-miR-486-5p	UCCUGUACUGAGCUGCCCCGAG
hsa-miR-23b-3p	AUCACAUUGCCAGGGAUUACC
hsa-miR-140-3p	UACCACAGGGUAGAACCACGG
hsa-miR-101-3p	UACAGUACUGUGAUAACUGAA
hsa-miR-107	AGCAGCAUUGUACAGGGCUAUCA
hsa-miR-130b-3p	CAGUGCAAUGAUGAAAGGGCAU
hsa-miR-369-3p	AAUAAUACAUGGUUGAUCUUU
hsa-miR-133a	UUUGGUCCCCUUCAACCAGCUG
hsa-miR-222-3p	AGCUACAUCUGGCUACUGGGU
hsa-miR-320d	AAAAGCUGGGUUGAGAGGA
hsa-miR-30a-5p	UGUAAACAUCCUCGACUGGAAG
hsa-miR-181a-5p	AACAUUCAACGCUGUCGGUGAGU
hsa-miR-140-5p	CAGUGGUUUUACCCUAUGGUAG
hsa-miR-425-3p	AUCGGGAAUGUCGUGUCCGCCC
hsa-miR-106b-3p	CCGCACUGUGGGUACUUGCUGC
hsa-miR-192-5p	CUGACCUAUGAAUUGACAGCC
hsa-miR-10a-3p	CAAAUUCGUAUCUAGGGGAAUA
hsa-miR-17-5p	CAAAGUGCUUACAGUGCAGGUAG
hsa-miR-590-5p	GAGCUUAUUCAUAAAAGUGCAG
hsa-miR-1299	UUCUGGAAUUCUGUGUGAGGGA
hsa-miR-365a-3p	UAAUGCCCCUAAAAAUCCUUAU
hsa-miR-500a-5p	UAAUCCUUGCUACCUGGGUGAGA
hsa-miR-32-5p	UAUUGCACAUUACUAAGUUGCA
hsa-miR-340-5p	UUAUAAAGCAAUGAGACUGAUU
hsa-miR-374b-5p	AUAUAAUACAACCUGCUAAGUG
hsa-miR-27a-3p	UUCACAGUGGCUAAGUUCCGC
hsa-miR-627	GUGAGUCUCUAAGAAAAGAGGA
hsa-miR-539-5p	GGAGAAAUUAUCCUUGGUGUGU
hsa-miR-342-5p	AGGGGUGCUAUCUGUGAUUGA
hsa-miR-484	UCAGGCUCAGUCCCCUCCCGAU
hsa-miR-132-3p	UAACAGUCUACAGCCAUGGUCG
hsa-miR-379-5p	UGGUAGACUAUGGAACGUAGG
hsa-miR-125a-3p	ACAGGUGAGGUUCUUGGGAGCC
hsa-miR-29a-3p	UAGCACCAUCUGAAAUCGGUUA

hsa-miR-363-3p	AAUUGCACGGUAUCCAUCUGUA
hsa-miR-376b	AUCAUAGAGGAAAAUCCAUGUU
hsa-miR-589-5p	UGAGAACCACGUCUGCUCUGAG
hsa-miR-432-5p	UCUUGGAGUAGGUCAUUGGGUGG
hsa-miR-1280	UCCCACCGCUGCCACCC
hsa-miR-103a-3p	AGCAGCAUUGUACAGGGCUAUGA
hsa-miR-122-5p	UGGAGUGUGACAAUGGUGUUUG
hsa-miR-93-5p	CAAAGUGCUGUUCGUGCAGGUAG
hsa-miR-25-3p	CAUUGCACUUGUCUCGGUCUGA
hsa-miR-9-5p	UCUUUGGUUAUCUAGCUGUAUGA
hsa-miR-579	UUCAUUUGGUAUAAACCGCGAUU
hsa-miR-136-3p	CAUCAUCGUCUCAAAUGAGUCU
hsa-miR-146a-5p	UGAGAACUGAAUUCCAUGGGUU
hsa-miR-144-5p	GGAUAUCAUCAUAUACUGUAAG
hsa-miR-15a-5p	UAGCAGCACAUAAUGGUUUGUG
hsa-miR-150-5p	UCUCCCAACCCUUGUACCAGUG
hsa-miR-152	UCAGUGCAUGACAGAACUUGG
hsa-miR-29c-5p	UGACCGAUUUCUCCUGGUGUUC
hsa-miR-320c	AAAAGCUGGGUUGAGAGGGU
hsa-miR-127-3p	UCGGAUCCGUCUGAGCUUGGCU
hsa-miR-331-5p	CUAGGUAUGGUCCCAGGGAUCC
hsa-miR-378a-3p	ACUGGACUUGGAGUCAGAAGG
hsa-miR-374a-5p	UUAUAAUACAACCUGAUAAGUG
hsa-miR-409-3p	GAAUGUUGCUCGGUGAACCCCU
hsa-miR-411-3p	UAUGUAACACGGUCCACUAACC
hsa-miR-505-3p	CGUCAACACUUGCUGGUUUCCU
hsa-miR-628-5p	AUGCUGACAUAUUUACUAGAGG
hsa-miR-629-3p	GUUCUCCCAACGUAAGCCCAGC
hsa-miR-4732-3p	GCCCUGACCUGUCCUGUUCUG
hsa-miR-501-5p	AAUCCUUUGUCCCUGGGUGAGA
hsa-miR-616-5p	ACUCAAAACCCUUCAGUGACUU
hsa-miR-454-3p	UAGUGCAAUAUUGCUUAUAGGGU
hsa-miR-485-3p	GUCAUACACGGCUCUCCUCUCU
hsa-miR-133b	UUUGGUCCCCUUCAACCAGCUA
hsa-miR-186-5p	CAAAGAAUUCUCCUUUUGGGCU
hsa-miR-20b-5p	CAAAGUGCUCAUAGUGCAGGUAG
hsa-miR-30d-5p	UGUAAACAUCCCCGACUGGAAG
hsa-miR-375	UUUGUUCGUUCGGCUCGCGUGA
hsa-miR-16-5p	UAGCAGCACGUAAAUAUUGGCG
hsa-miR-106b-5p	UAAAGUGCUGACAGUGCAGAU
hsa-miR-139-5p	UCUACAGUGCACGUGUCUCCAG

hsa-miR-141-3p	UAACACUGUCUGGUAAAGAUGG
hsa-miR-185-5p	UGGAGAGAAAGGCAGUUCCUGA
hsa-miR-181b-5p	AACAUUCAUUGCUGUCGGUGGGU
hsa-miR-199a-3p	ACAGUAGUCUGCACAUUGGUUA
hsa-miR-19b-3p	UGUGCAAAUCCAUGCAAAACUGA
hsa-miR-148b-3p	UCAGUGCAUCACAGAACUUUGU
hsa-miR-29b-3p	UAGCACCAUUUGAAAUCAGUGUU
hsa-miR-338-5p	AACAAUAUCCUGGUGCUGAGUG
hsa-miR-584-5p	UUAUGGUUUGCCUGGGACUGAG
hsa-miR-382-5p	GAAGUUGUUCGUGGUGGAUUCG
hsa-miR-151a-3p	CUAGACUGAAGCUCCUUGAGG
hsa-miR-1290	UGGAUUUUUGGAUCAGGGA
hsa-miR-200b-3p	UAAUACUGCCUGGUAAUGAUGA
hsa-miR-411-5p	UAGUAGACCGUAUAGCGUACG
hsa-miR-126-5p	CAUUAUUACUUUUGGUACGCG
hsa-miR-101-5p	CAGUUAUCACAGUGCUGAUGCU
hsa-miR-125b-5p	UCCCUGAGACCCUAACUUGUGA
hsa-miR-362-5p	AAUCCUUGGAACCUAGGUGUGAGU
hsa-miR-197-3p	UUCACCACCUUCUCCACCCAGC
hsa-miR-221-3p	AGCUACAUUGUCUGCUGGGUUUC
hsa-miR-501-3p	AAUGCACCCGGGCAAGGAUUCU
hsa-miR-671-3p	UCCGGUUCUCAGGGCUCCACC
hsa-miR-181a-2-3p	ACCACUGACCGUUGACUGUACC
hsa-miR-9-3p	AUAAAGCUAGAUAACCGAAAGU
hsa-miR-452-5p	AACUGUUUGCAGAGGAAACUGA
hsa-miR-598	UACGUCAUCGUUGUCAUCGUCA
hsa-miR-320b	AAAAGCUGGGUUGAGAGGGCAA
hsa-miR-328	CUGGCCCUCUCUGCCCUUCCGU
hsa-miR-650	AGGAGGCAGCGCUCUCAGGAC
hsa-miR-134	UGUGACUGGUUGACCAGAGGGG
hsa-miR-130a-3p	CAGUGCAAUGUUAAAAGGGCAU
hsa-miR-21-5p	UAGCUUAUCAGACUGAUGUUGA
hsa-miR-424-5p	CAGCAGCAAUUCAUGUUUUGAA
hsa-miR-99a-5p	AACCCGUAGAUCCGAUCUUGUG
hsa-miR-18a-3p	ACUGCCCUAAGUGCUCCUUCUGG
hsa-miR-195-5p	UAGCAGCACAGAAAUAUUGGC
hsa-miR-205-5p	UCCUUCAUUCCACCGGAGUCUG
hsa-miR-206	UGGAAUGUAAGGAAGUGUGUGG
hsa-miR-500a-3p	AUGCACCUGGGCAAGGAUUCUG
hsa-miR-18b-5p	UAAGGUGCAUCUAGUGCAGUUAG
hsa-miR-181d	AACAUUCAUUGUUGUCGGUGGGU

hsa-miR-339-3p	UGAGCGCCUCGACGACAGAGCCG
hsa-miR-93-3p	ACUGCUGAGCUAGCACUUCCCG
hsa-miR-10b-5p	UACCCUGUAGAACCGAAUUUGUG
hsa-miR-497-5p	CAGCAGCACACUGUGGUUUGU
hsa-miR-27b-3p	UUCACAGUGGCUAAGUUCUGC
hsa-miR-128	UCACAGUGAACCGGUCUCUUU
hsa-miR-183-5p	UAUGGCACUGGUAGAAUUCACU
hsa-miR-22-3p	AAGCUGCCAGUUGAAGAACUGU
hsa-miR-26a-5p	UUCAAGUAAUCCAGGAUAGGCU
hsa-miR-223-3p	UGUCAGUUUGUCAAAUACCCCA
hsa-miR-629-5p	UGGGUUUACGUUGGGAGAACU
hsa-miR-92a-3p	UAUUGCACUUGUCCCGGCCUGU
hsa-miR-29b-2-5p	CUGGUUUCACAUGGUGGCUUAG
hsa-miR-21-3p	CAACACCAGUCGAUGGGCUGU
hsa-miR-199a-5p	CCCAGUGUUCAGACUACCUGUUC
hsa-miR-148a-3p	UCAGUGCACUACAGAACUUUGU
hsa-miR-193a-5p	UGGGUCUUUGCGGGCGAGAUGA
hsa-miR-27a-5p	AGGGCUUAGCUGCUUGUGAGCA
hsa-miR-200c-3p	UAAUACUGCCGGGUAAUGAUGGA
hsa-miR-20a-5p	UAAAGUGCUUAUAGUGCAGGUAG
hsa-miR-194-5p	UGUAACAGCAACUCCAUGUGGA
hsa-miR-532-3p	CCUCCCACACCCAAGGCUUGCA
hsa-miR-19a-3p	UGUGCAAAUCUAUGCAAAACUGA
hsa-miR-142-5p	CAUAAAGUAGAAAGCACUACU
hsa-miR-144-3p	UACAGUAUAGAUGAUGUACU
hsa-miR-145-5p	GUCCAGUUUUCCCAGGAAUCCCU
hsa-miR-10a-5p	UACCCUGUAGAUCCGAAUUUGUG
hsa-miR-23a-3p	AUCACAUUGCCAGGGAUUUCC
hsa-miR-23a-5p	GGGGUUCCUGGGGAUGGGAUUU
hsa-miR-15b-3p	CGAAUCAUUAUUUGCUGCUCUA
hsa-miR-301a-3p	CAGUGCAAUAGUAUUGUCAAAGC
hsa-miR-660-5p	UACCCAUUGCAUAUCGGAGUUG
hsa-miR-30b-5p	UGUAAACAUCCUACACUCAGCU
hsa-miR-30e-5p	UGUAAACAUCCUUGACUGGAAG
hsa-miR-550a-5p	AGUGCCUGAGGGAGUAAGAGCCC
hsa-miR-425-5p	AAUGACACGAUCACUCCCGUUGA
hsa-miR-4306	UGGAGAGAAAGGCAGUA
hsa-miR-532-5p	CAUGCCUUGAGUGUAGGACCGU
hsa-miR-335-5p	UCAAGAGCAAUAACGAAAAAUGU
hsa-miR-483-5p	AAGACGGGAGGAAAGAAGGGAG
hsa-miR-1226-3p	UCACCAGCCCUGUGUUCCCUAG

hsa-miR-431-5p	UGUCUUGCAGGCCGUCAUGCA
hsa-miR-324-5p	CGCAUCCCCUAGGGCAUUGGUGU
hsa-miR-487b	AAUCGUACAGGGUCAUCCACUU
hsa-miR-451a	AAACCGUUACCAUUACUGAGUU
hsa-miR-493-5p	UUGUACAUGGUAGGCUUUCAUU
hsa-miR-136-5p	ACUCCAUUUGUUUUGAUGAUGGA
hsa-miR-23c	AUCACAUUGCCAGUGAUUACCC
hsa-miR-95	UUCAACGGGUAUUUAUUGAGCA
hsa-miR-423-5p	UGAGGGGCAGAGAGCGAGACUUU
hsa-miR-320e	AAAGCUGGGUUGAGAAGG
hsa-miR-224-5p	CAAGUCACUAGUGGUUCCGUU
hsa-miR-28-3p	CACUAGAUUGUGAGCUCCUGGA
hsa-miR-29c-3p	UAGCACCAUUUGAAAUCGGUUA
hsa-miR-326	CCUCUGGGCCCUUCCUCCAG
hsa-miR-596	AAGCCUGCCCGGCUCCUCGGG
hsa-miR-885-5p	UCCAUUACACUACCCUGCCUCU
hsa-miR-146b-5p	UGAGAACUGAAUUCCAUAGGCU
hsa-miR-34a-5p	UGGCAGUGUCUUAGCUGGUUGU
hsa-miR-330-3p	GCAAAGCACACGGCCUGCAGAGA
hsa-miR-154-5p	UAGGUUAUCCGUGUUGCCUUCG
hsa-miR-191-5p	CAACGGAAUCCCAAAAGCAGCUG
hsa-miR-193b-3p	AACUGGCCCUCAAAGUCCCGCU
hsa-miR-301b	CAGUGCAAUGAUAUUGUCAAAGC
hsa-miR-30e-3p	CUUUCAGUCGGAUGUUUACAGC
hsa-miR-320a	AAAAGCUGGGUUGAGAGGGCGA
hsa-miR-199b-3p	ACAGUAGUCUGCACAUUGGUUA
hsa-miR-502-3p	AAUGCACCUGGGCAAGGAUUCA
hsa-miR-450a-5p	UUUUGCGAUGUGUUCCUAAUAU
hsa-miR-495	AAACAAACAUGGUGCACUUCUU
hsa-miR-126-3p	UCGUACCGUGAGUAAUAAUGCG
hsa-miR-15b-5p	UAGCAGCACAUCAUGGUUUACA
hsa-miR-339-5p	UCCCUGUCCUCCAGGAGCUCACG
hsa-miR-337-5p	GAACGGCUUCAUACAGGAGUU

Table S2. MiRNAs Differentially Expressed between Normal and Gastric Cancer. For the comparison between normal and all gastric cancer subjects (regardless of subtypes and stages), 75 miRNA had p-value lower than 0.01 after FDR correction (Bonferroni method). AUC – area under the receiver operating characteristic curve; fold change – the mean expression level (copy/ml) of miRNA in the cancer population divided by that in the normal population.

Up-regulated miRNAs

miRNA name	AUC	P-value	P-value, FDR correction	Fold change	Novel Observation
miR-101-3p	0.61	1.80E-05	9.50E-05	1.27	Novel
miR-106b-3p	0.66	3.70E-09	4.10E-08	1.13	Novel
miR-106b-5p	0.61	5.10E-04	1.70E-03	1.21	
miR-128	0.62	1.40E-06	8.80E-06	1.16	Novel
miR-1280	0.66	3.10E-09	3.90E-08	1.38	Novel
miR-140-3p	0.62	6.40E-06	3.50E-05	1.2	Novel
miR-140-5p	0.67	6.20E-10	1.20E-08	1.24	Novel
miR-142-5p	0.71	1.90E-14	3.70E-12	1.31	Novel
miR-148a-3p	0.67	2.20E-10	4.80E-09	1.32	Novel
miR-15b-3p	0.62	6.20E-06	3.50E-05	1.32	Novel
miR-17-5p	0.63	1.00E-05	5.50E-05	1.24	
miR-183-5p	0.64	8.80E-07	6.20E-06	1.53	Novel
miR-186-5p	0.59	1.40E-03	3.70E-03	1.11	Novel
miR-18b-5p	0.64	1.50E-07	1.20E-06	1.38	Novel
miR-197-3p	0.68	8.10E-13	5.10E-11	1.32	Novel
miR-19a-3p	0.63	4.90E-07	3.60E-06	1.29	Novel
miR-19b-3p	0.59	1.10E-03	3.00E-03	1.18	Novel
miR-20a-5p	0.65	1.20E-07	1.10E-06	1.35	
miR-20b-5p	0.60	2.90E-04	9.90E-04	1.3	Novel
miR-21-3p	0.60	7.90E-05	3.20E-04	1.13	Novel
miR-21-5p	0.63	2.60E-08	2.80E-07	1.23	
miR-223-3p	0.66	7.00E-10	1.20E-08	1.36	
miR-23a-5p	0.64	1.00E-07	9.20E-07	1.31	Novel
miR-25-3p	0.62	3.40E-05	1.60E-04	1.26	Novel
miR-27a-5p	0.69	1.00E-13	1.00E-11	1.76	
miR-29a-3p	0.61	6.00E-05	2.60E-04	1.17	Novel
miR-29b-2-5p	0.59	4.70E-05	2.10E-04	1.16	Novel
miR-29b-3p	0.61	7.20E-05	3.00E-04	1.18	Novel
miR-29c-3p	0.65	2.00E-09	2.90E-08	1.23	Novel
miR-29c-5p	0.63	1.40E-06	8.80E-06	1.15	Novel
miR-338-5p	0.57	3.70E-03	9.40E-03	1.29	Novel
miR-423-5p	0.60	7.20E-05	3.00E-04	1.18	
miR-424-5p	0.68	7.00E-11	1.90E-09	1.41	Novel

miR-425-3p	0.57	2.20E-03	5.70E-03	1.05	Novel
miR-4306	0.63	1.20E-06	8.00E-06	1.35	Novel
miR-450a-5p	0.67	2.10E-10	4.80E-09	1.53	Novel
miR-486-5p	0.61	9.60E-05	3.70E-04	1.32	Novel
miR-500a-3p	0.60	1.10E-04	4.20E-04	1.2	Novel
miR-501-5p	0.60	9.60E-04	2.80E-03	1.24	Novel
miR-532-3p	0.60	1.90E-04	7.00E-04	1.15	Novel
miR-550a-5p	0.63	9.00E-07	6.20E-06	1.38	Novel
miR-579	0.62	2.20E-05	1.10E-04	1.3	Novel
miR-589-5p	0.63	1.70E-06	1.00E-05	1.18	Novel
miR-590-5p	0.69	3.00E-12	1.40E-10	1.23	Novel
miR-598	0.67	7.10E-12	2.70E-10	1.27	Novel
miR-616-5p	0.65	3.40E-09	4.10E-08	1.35	Novel
miR-627	0.58	7.30E-04	2.30E-03	1.19	Novel
miR-629-3p	0.67	6.10E-11	1.90E-09	1.38	Novel
miR-629-5p	0.63	1.40E-04	5.10E-04	1.5	Novel
miR-93-3p	0.62	5.10E-06	3.00E-05	1.22	Novel
miR-93-5p	0.60	2.30E-04	8.00E-04	1.21	Novel

Down-regulated miRNAs

miRNA name	AUC	P-value	P-value, FDR correction	Fold change	
miR-107	0.65	4.40E-08	4.40E-07	0.8	Novel
miR-122-5p	0.61	8.10E-05	3.20E-04	0.66	Novel
miR-126-3p	0.66	1.70E-09	2.70E-08	0.87	Novel
miR-136-5p	0.61	2.30E-05	1.10E-04	0.72	Novel
miR-139-5p	0.60	8.60E-05	3.40E-04	0.84	Novel
miR-146a-5p	0.59	2.10E-03	5.60E-03	0.89	Novel
miR-154-5p	0.59	8.60E-04	2.60E-03	0.8	Novel
miR-181a-5p	0.60	2.30E-04	8.00E-04	0.92	Novel
miR-193b-3p	0.58	1.20E-03	3.20E-03	0.77	Novel
miR-23c	0.59	8.00E-04	2.40E-03	0.84	Novel
miR-26a-5p	0.60	4.40E-05	2.00E-04	0.86	Novel
miR-30a-5p	0.64	6.70E-08	6.40E-07	0.76	Novel
miR-30b-5p	0.59	9.50E-04	2.80E-03	0.9	Novel
miR-337-5p	0.63	4.80E-07	3.60E-06	0.74	Novel
miR-339-5p	0.64	4.90E-07	3.60E-06	0.79	Novel
miR-382-5p	0.59	1.00E-03	2.90E-03	0.81	Novel

miR-409-3p	0.59	5.00E-04	1.60E-03	0.77	Novel
miR-411-5p	0.6	7.30E-04	2.30E-03	0.74	Novel
miR-485-3p	0.6	6.40E-04	2.00E-03	0.77	Novel
miR-487b	0.59	1.10E-03	3.00E-03	0.76	Novel
miR-495	0.6	2.10E-04	7.40E-04	0.77	Novel
miR-885-5p	0.62	1.90E-05	9.60E-05	0.69	Novel
miR-99a-5p	0.58	2.90E-03	7.50E-03	0.82	Novel
miR-99b-5p	0.67	2.60E-09	3.50E-08	0.78	Novel

Table S3. Summary of Serum / Plasma miRNA Biomarker Studies for Gastric Cancer. The studies that measured the cell-free serum/plasma miRNAs were included in the table. Only the results validated with RT-qPCR were shown. GC: gastric cancer subjects. C: control subjects.

Paper	Up regulated	Down regulated	Method	Samples
Chen Li et al [13]	miR-199a-3p	-	RT-qPCR	Plasma/80GC/70C
Aysegul Gorur et al [14]	-	miR-195-5p	RT-qPCR	Serum/20GC/190C
Hui Cai et al [15]	miR-106b, miR-20a, miR-221	-	RT-qPCR	Plasma/90GC/90C
Mei-Hua Cui et al [16]	miR-181c	-	RT-qPCR	Plasma/30GC/60C
Chen Li et al [17]	miR-199a-3p, miR-151-5p	-	RT-qPCR	Plasma/180GC/100C
Ming-yang Song et al [18]	miR-221, miR-744, miR-376c, miR-191, miR-27a, let-7e, miR-27b, and miR-222	-	RT-qPCR	Serum/82GC/82C
Bo-sheng Li et al [19]	miR-223, miR-21	miR- 218	RT-qPCR	Plasma/60GC/60C
Manuel Valladares-Ayerbes et al [20]	miR-200c	-	RT-qPCR	whole blood/52GC/15C
Wen-Hui Zhang et al [21]	-	miR-375	RT-qPCR	Serum
S. S. Lo et al [22]	miR-370	-	RT-qPCR	Plasma/33GC/33C
M. Tsujiura et al [23]	miR-17-5p, miR-21, miR-106a, miR-106b	let-7a	RT-qPCR	Plasma/69GC/30C
Rui Liu et al [24]	miR-1, miR-20a, miR-27a, miR-34a, miR-423-5p	-	RT-qPCR	Serum/142GC/105C
Hanshao Liu et al [25]	miR-187*,miR-371-5p, miR-378	-	RT-qPCR	Serum/40GC/41C

Table S4. MiRNAs Differentially Expressed between Different Stages of Gastric Cancer. A total of 36 miRNAs with p-value lower than 0.05 were identified from the comparison mad with the four stages of gastric cancer, based on two-way anova test (subtypes and stages) after false discovery rate correction (Bonferroni method). The expression levels (copy/ml) were analyzed based on the log2 scale. For each miRNA, the significant levels for the alternations between stage 1 and 2, stage 2 and 3, stage 3 and 4 were calculated based on anova test and Bonferroni adjustment to compensate for multiple comparisons. *: p-value < 0.05; **: p-value < 0.01; ***: p-value < 0.001. A miRNA was considered up-regulated if its expression level was higher in the later stage.

	anova p<0.01	Group	change between stage 2 and stage 1	change between stage 3 and stage 2	change between stage 4 and stage 3	change between normal and all cancer
hsa-miR-27a-3p	0.0008	I	-	down**	up***	No change
hsa-miR-1280	0.0007	C	up***	-	-	up-regulated
hsa-miR-579	0.0067	A	-	-	-	up-regulated
hsa-miR-150-5p	0.0035	F	-	-	-	No change
hsa-miR-29c-5p	0.0010	B	up***	-	-	up-regulated
hsa-miR-186-5p	0.0087	H	up**	down**	-	up-regulated
hsa-miR-338-5p	0.0072	L	-	-	up**	up-regulated
hsa-miR-362-5p	0.0014	B	up**	-	-	No change
hsa-miR-197-3p	0.0000	B	up***	-	-	up-regulated
hsa-miR-221-3p	0.0000	B	up***	-	-	No change
hsa-miR-501-3p	0.0000	C	up***	-	down*	No change
hsa-miR-181a-2-3p	0.0072	G	up**	-	-	No change
hsa-miR-598	0.0000	A	up**	-	-	up-regulated
hsa-miR-320b	0.0014	C	up**	-	-	No change
hsa-miR-328	0.0000	C	up***	-	down*	No change
hsa-miR-134	0.0072	D	-	-	down***	No change
hsa-miR-21-5p	0.0000	E	-	up**	down**	up-regulated
hsa-miR-424-5p	0.0000	B	up***	-	-	up-regulated
hsa-miR-99a-5p	0.0023	G	up*	-	down**	down-regulated
hsa-miR-18a-3p	0.0016	B	up***	-	-	No change
hsa-miR-195-5p	0.0000	G	up***	-	down**	No change
hsa-miR-500a-3p	0.0000	C	up***	-	down**	up-regulated
hsa-miR-18b-5p	0.0072	C	up*	-	-	up-regulated
hsa-miR-339-3p	0.0005	C	up***	-	-	No change
hsa-miR-128	0.0000	C	up***	-	down***	up-regulated
hsa-miR-22-3p	0.0016	C	-	-	down**	No change
hsa-miR-26a-5p	0.0002	G	up**	-	down***	down-regulated
hsa-miR-29b-2-5p	0.0087	B	up*	-	-	up-regulated

hsa-miR-148a-3p	0.0029	A	-	-	-	up-regulated
hsa-miR-142-5p	0.0004	H	up**	down***	up**	up-regulated
hsa-miR-23a-3p	0.0000	B	up***	-	-	No change
hsa-miR-23c	0.0002	C	up**	-	down**	down-regulated
hsa-miR-28-3p	0.0072	K	down*	-	-	No change
hsa-miR-193b-3p	0.0029	K	down**	-	-	down-regulated
hsa-miR-320a	0.0004	J	down***	up**	-	No change
hsa-miR-15b-5p	0.0000	B	up***	down*	-	No change

Table S5. Multivariate Cox regression analysis of 12-miRNA panel and clinical covariates

Variables		**Log Hazard Ratio, ln(HR)**	***p*-value**
12-miR	Cancer vs Non-cancer	13.9	< 0.001
Age (years)	> 50 vs ≤ 50	0.66	0.04
Gender	Male vs Female	0.54	0.47
Ethnicity	Chinese vs Non-Chinese	0.17	0.64
H. pylori	Yes vs No	-0.04	0.89

Table S6. Cross-reactivity test against other common cancers

#	Type of Cancer	Number of specimens tested	Number of specimen with high risk score based on GASTROClear
1	Esophageal	12	1
2	Liver	6	1
3	Colorectal	12	3
4	Lung	12	1
5	Breast	12	0
6	Prostate	12	0
7	Kidney	12	5
8	Bladder	12	0
	Total	**90**	**11**

So JBY, *et al. Gut* 2020;0:1–9. doi: 10.1136/gutjnl-2020-322065

Table S7. Base-case Values and Corresponding Sensitivity Range for Variables in Cost-Effectiveness Modelling

Singaporean healthcare setup			
Variable name	**Base-case value**	**Sensitivity Range**	**Source**
Costs (USD)			
MiRNA test	30	10 – 500	Assumed
Upper -endoscopy (EGD)	493	100 - 500	National University Hospital, Singapore (NUH)
Biopsy	122	-	
Stage 1 treatment	10423	-	
Stage 2 treatment	10423	-	
Stage 3 treatment	29451	-	
Stage 4 treatment	3069	-	
Follow-up examinations	719	-	
Staging Investigation (EUS + CT+ CXR+ follow-up)	1513	-	
Probabilities			
Incidence of Gastric Cancer in Chinese Males by Age group			Report No.8, 2015. Singapore Cancer Registry [26]
50 - 54 years	0.018%		
55 - 59 years	0.029%		
60 - 64 years	0.053%		
65 - 69 years	0.098%		
70 - 74 years	0.157%		
75 years	0.187%		
Stage specific diagnosis currently Stage 1: 2 : 3 : 4	18% : 11.5% : 27.5% : 43%		
Recurrence of Gastric Cancer in successfully treated patients by stage			
Recurrence in Stage 1 patients	11%	5% - 30%	Roukos *et al.* [6]
Recurrence in Stage 2 patients	53%	30% - 60%	
Recurrence in Stage 3 patients	83%	50% - 90%	

Utility Values (disutility*)			
Stage 1	0.88 (0.28)	0.78 – 0.98	Zhou HJ *et al.* [11]
Stage 2	0.86 (0.29)	0.72 – 0.99	
Stage 3	0.77 (0.31)	0.57 – 0.97	
Stage 4	0.68 (0.08)	0.52 – 0.84	
Test Characteristics			
Endoscopy Sensitivity	93%	-	Voutilainen *et al.* [12] Hamashima et *al. [27]*
Endoscopy Specificity	100%	-	Voutilainen *et al.* [12]
miRNA Sensitivity by Stages (Stage 1:2:3:4)	63% : 75% : 89% : 93%	30% - 100%	Current Study
miRNA specificity	89%	60% - 100%	Current Study

*Disutility refers to temporary reduction in QoL during first 6 months of treatment. Note: Assumed treatments are based on observed practice in Singapore. Gastric cancer patient on diagnosis undergoes staging investigation (CT, CXR, EUS & specialist consultation). Curative treatment includes surgery (total/ partial gastrectomy) & hospital stay (12days). Stage 3 patients undergo additional chemo-radiotherapy. Follow ups include: visits (2.2/year), repeat CT, CXR (1.4/year). Palliative care includes bypass surgery (30%), endoscopic stenting (6%), palliative chemotherapy (16%) & conservative treatment (2x specialist visits) (48%) with an appropriate hospital stay (12 days - on surgery, 2.5 days on average - if no surgery is performed). *Abbreviations used: CT: Computerized Tomography; CXR: Chest X-Ray; EUS: Endoscopic Ultrasound*

REFERENCE

1 Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A*, et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002;**3**:RESEARCH0034.

2 Andersen CL, Jensen JL, Orntoft TF. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res 2004;**64**:5245-50.

3 Singapore Cancer Registry Advisory Committe and Health Promotion Board S. Trends in Cancer Incidence in Singapore 2009-2013. 2014.

4 Singapore Cancer Registry. Cancer incidence and mortality 2003 - 2012 and selected trends 1973 - 2012 in Singapore. National Registry of Disease Office, 2015.

5 Department of Statistics MoTaI, Republic of Singapore. Population Trends 2016. 2016.

6 Roukos DH, Lorenz M, Karakostas K, Paraschou P, Batsis C, Kappas AM. Pathological serosa and node-based classification accurately predicts gastric cancer recurrence risk and outcome, and determines potential and limitation of a Japanese-style extensive surgery for Western patients: a prospective with quality control 10-year follow-up study. Br J Cancer 2001;**84**:1602-9.

7 Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol 2005;**5**:13.

8 Dan YY, So JB, Yeoh KG. Endoscopic screening for gastric cancer. Clin Gastroenterol Hepatol 2006;**4**:709-16.

9 Zhou HJ, Dan YY, Naidoo N, Li SC, Yeoh KG. A cost-effectiveness analysis evaluating endoscopic surveillance for gastric cancer for populations with low to intermediate risk. PLoS ONE 2013;**8**:e83959.

10 Lee S, Jun JK, Suh M, Park B, Noh DK, Jung K-W*, et al.* Gastric Cancer Screening Uptake Trends in Korea: Results for the National Cancer Screening Program From 2002 to 2011: A Prospective Cross-Sectional Study. Medicine 2015;**94**:e533.

11 Zhou HJ, So JB, Yong WP, Luo N, Zhu F, Naidoo N*, et al.* Validation of the functional assessment of cancer therapy-gastric module for the Chinese population. Health and quality of life outcomes 2012;**10**:145.

12 Voutilainen ME, Juhola MT. Evaluation of the diagnostic accuracy of gastroscopy to detect gastric tumours: clinicopathological features and prognosis of patients with gastric cancer missed on endoscopy. European journal of gastroenterology & hepatology 2005;**17**:1345-9.

13 Li C, Li J, Cai Q, Qiu Q, Yan M, Liu B*, et al.* MiRNA-199a-3p: A potential circulating diagnostic biomarker for early gastric cancer. Journal of surgical oncology 2013;**108**:89-92.

14 Gorur A, Balci Fidanci S, Dogruer Unal N, Ayaz L, Akbayir S, Yildirim Yaroglu H*, et al.* Determination of plasma microRNA for early detection of gastric cancer. Molecular biology reports 2013;**40**:2091-6.

15 Cai H, Yuan Y, Hao Y-F, Guo T-K, Wei X, Zhang Y-M. Plasma microRNAs serve as novel potential biomarkers for early detection of gastric cancer. Medical oncology (Northwood, London, England) 2013;**30**:452.

16 Cui M-H, Hou X-L, Lei X-Y, Mu F-H, Yang G-B, Yue L*, et al.* Upregulation of microRNA 181c expression in gastric cancer tissues and plasma. Asian Pacific journal of cancer prevention : APJCP 2013;**14**:3063-6.

17 Li C, Li J, Cai Q, Qiu Q, Yan M, Liu B*, et al.* miRNA-199a-3p in plasma as a potential diagnostic biomarker for gastric cancer. Annals of surgical oncology 2013;**20 Suppl 3**:405.

18 Song M-y, Pan K-f, Su H-j, Zhang L, Ma J-l, Li J-y*, et al.* Identification of serum microRNAs as novel non-invasive biomarkers for early detection of gastric cancer. PloS one 2012;**7**.

19 Li B-s, Zhao Y-l, Guo G, Li W, Zhu E-d, Luo X*, et al.* Plasma microRNAs, miR-223, miR-21 and miR-218, as novel potential biomarkers for gastric cancer detection. PloS one 2012;**7**.
20 Valladares-Ayerbes M, Reboredo M, Medina-Villaamil V, Iglesias-Díaz P, Lorenzo-Patiño M, Haz M*, et al.* Circulating miR-200c as a diagnostic and prognostic biomarker for gastric cancer. Journal of translational medicine 2012;**10**:186.
21 Zhang W-H, Gui J-H, Wang C-Z, Chang Q, Xu S-P, Cai C-H*, et al.* The identification of miR-375 as a potential biomarker in distal gastric adenocarcinoma. Oncology research 2012;**20**:139-47.
22 Lo SS, Hung PS, Chen JH, Tu HF, Fang WL, Chen CY*, et al.* Overexpression of miR-370 and downregulation of its novel target TGFβ-RII contribute to the progression of gastric carcinoma. Oncogene 2012;**31**:226-37.
23 Tsujiura M, Ichikawa D, Komatsu S, Shiozaki A, Takeshita H, Kosuga T*, et al.* Circulating microRNAs in plasma of patients with gastric cancers. Br J Cancer 2010;**102**:1174-9.
24 Liu R, Zhang C, Hu Z, Li G, Wang C, Yang C*, et al.* A five-microRNA signature identified from genome-wide serum microRNA expression profiling serves as a fingerprint for gastric cancer diagnosis. European journal of cancer (Oxford, England : 1990) 2011;**47**:784-91.
25 Liu H, Zhu L, Liu B, Yang L, Meng X, Zhang W*, et al.* Genome-wide microRNA profiles identify miR-378 as a serum biomarker for early detection of gastric cancer. Cancer Lett 2012;**316**:196-203.
26 Office NRoD. Cancer Incidence and Mortality 2003-12 and Selected Trends 1973-2012 In Singapore. Singapore Cancer Registry, Report No. 8. 2015.
27 Hamashima C, Okamoto M, Shabana M, Osaki Y, Kishimoto T. Sensitivity of endoscopic screening for gastric cancer by the incidence method. International journal of cancer Journal international du cancer 2013;**133**:653-9.

Supplementary Material

Supplementary Methods

Analytical Validation of miRNA Assays

The analytical performance of the miRNA assays used for gastric cancer biomarker identification was evaluated. We first evaluated the analytical specificity of these assays by conducting a cross-reactivity test of miRNA assays against 9 highly homologous let-7 family members (**Figure S2A**), a design routinely used to evaluate assay specificity. The let-7 family assays were able to discriminate homologous sequences with even single nucleotide differences, e.g. let-7a assay showed 100% detection of let-7a target and only 0.8% cross-reactivity against let-7c target. Secondly, we evaluated the reproducibility of the miRNA assays by measuring 200 circulating miRNAs in 30 control and cancer serum specimen in two independent labs (**Figure S2B**). After normalization of technical and biological variations using the multi-layered controls illustrated in **Figure S1**, these assays demonstrated encouraging concordance of 0.95-0.98 in all 30 clinical samples. Lastly, we evaluated the analytical sensitivity of the miRNA assays (**Figure S2C**). Constrained by the small size, miRNA assay performances can be highly variable across different miRNA targets, especially miRNAs with higher AT content. We selected 8 commonly studied miRNAs with low to high AT content (36.4% - 63.6%) and compared the analytical sensitivity and dynamic range of the MiRXES miRNA assays against the well known Taqman probe based assays. The miRNA assays used for gastric cancer biomarker discovery demonstrated consistent amplification and detection of all 8 miRNA targets across at least 7 logs of dynamic range where the probe-based assays showed less consistent performance, especially against miRNA targets with higher AT content. Overall, these validation studies demonstrated good analytical performance of the assays and the workflow, and warrant their use for biomarker discovery.

Laboratory Procedures for miRNA Expression Quantification in Discovery and Verification Phases

Spike-In Controls for RT-qPCR Workflow

To monitor and normalize technical variations in RNA isolation efficiency, a set of 3 proprietary synthetic miRNAs were spiked into the sample lysis buffer (Qiazol) at high, medium and low concentrations. To monitor and normalize technical variations in subsequent RT and qPCR reactions, a second set of 3 proprietary synthetic miRNAs were then spiked into each isolated sample RNA at high, medium and low concentrations. A 6-log serial dilution of synthetic templates (10^7 to 10^2 copies) for each miRNA, non-template control (nuclease-free water spiked with MS2) and reference human serum RNA were concurrently reversed-transcribed and quantified by qPCR with each isolated serum RNA sample. These control measures facilitated monitoring and normalization of technical variations in pipetting and assay efficiency in RT, cDNA amplification, and qPCR.

Determination and Normalization of Absolute miRNA Expressions

Upon completion of RT-qPCR, Ct values were determined using the ViiA 7 RUO software (Thermo Fisher Scientific Inc, USA) with automatic baseline setting and a threshold of 0.5. Absolute expression of each miRNA in patient serum was determined through intra-polation of synthetic miRNA standard curves and corrected for RT-qPCR efficiency variation using spike-in RNAs. The miRNA expression of each sample was further normalized using 6 endogenous reference miRNAs independently identified using the geNorm and NormFinder reference gene algorithms [1, 2]. The miRNA expression profiles normalized using the 6 reference miRNAs were similar to that normalized by global mean expression of all miRNA quantified. Absolute expression of miRNAs were log2 transformed for subsequent statistical analysis and optimization of multivariate biomarker panel.

Multivariate Analysis For Constructing Multi-miR Panels

A linear support vector machine (SVM) was used to construct the multi-variant biomarker panels and the associated algorithm that classified cancer and control groups with highest AUC. Multiple iterations of four-

So JBY, *et al. Gut* 2021; 70:829–837. doi: 10.1136/gutjnl-2020-322065

fold cross-validation (matched by sex, cancer subtype and disease stage) were conducted to evaluate the performance of these panels. All calculations were performed using Matlab® software (MathWorks, USA).

Laboratory Procedures for 12-miR Multi-Target Assay in Validation Phase

The assay involved 3 steps: (1) RNA isolation from serum samples; (2) cDNA synthesis; and (3) Detection of miRNAs by quantitative PCR (qPCR). Extraction of RNA was performed by combining phenol/guanidine-based lysis of serum sample and silica-membrane-based purification of total RNA. During cDNA synthesis, 12 miRNA targets from each specimen were converted into cDNAs using 12 corresponding miRNA-specific stem-loop-based reverse transcription primers in a single reaction on a Veriti Dx thermocycler (Thermo Fisher Scientific Inc, USA). At the qPCR step, each miRNA target was amplified by a sequence-specific forward PCR primer and a hemi-nested sequence specific reverse PCR primer and detected using SYBR Green I dye in single-plex reactions on a Quantstudio Dx (384-well) real-time qPCR instrument (Thermo Fisher Scientific Inc, USA). Ct values of the 12 biomarker and reference miRNAs were exported using the QuantStudio Dx Software v1.0.1 (Thermo Fisher Scientific Inc, USA) and converted into a single numerical score using a validated, prespecified logistic-regression algorithm through the GASTROSmart Software (MIRXES Pte Ltd, Singapore). In each assay run, 13 patient specimens were processed concurrently with 2 quantitative reference specimen and 1 negative control specimen, which served as quality control and inter-run normalizers.

Cost-effectiveness analysis

We examined the cost-effectiveness of implementing the miRNA biomarker panel as a screen before endoscopy in a proposed national screening program in Singapore. Our study focusses on the cohort of Singaporean Chinese males, age 50-75 years, who are at an intermediate risk of gastric cancer, and compare the proposed mass screening program with the current pattern of gastric cancer diagnosis without screening. Chinese population carry ~90% of gastric cancer disease burden in Singapore with males at a 30% higher risk of gastric cancer than females [3, 4]. With the cancer incidence rising sharply after the age of 50 years[4], this subgroup with intermediate gastric cancer risk has a 4 times higher annual incidence rate than the general population. We estimate the quality-adjusted life years (QALY), costs per person, incremental cost-effectiveness ratio (ICER) and also the benefits of early cancer diagnosis and reduced mortality achieved by implementing the mass screening program.

Detailed research methodology of cost-effectiveness analysis is as follows:

- **Target Population:** The analysis is performed on the cohort of Singaporean Chinese males aged 50-75 years.
- **Interventions Compared**
  The two interventions compared are:
  1. Current practice of no screening.
  2. Mass screening program using miRNA -test, followed with test-positive patients undergoing a confirmatory upper-endoscopy and biopsy and test-negative subjects to be followed up 3-yearly.

- **Methodology**
  Markov decision model was built in Microsoft Excel 2010 to compare the two interventions in the target population by analyzing in a closed cohort setting (**Figure S7**). Model was populated using local and published data with the cohort size estimated from 2016 population census [5]. With a healthcare system perspective, a 25 years' time horizon was analyzed with subjects exiting the model at the age of 75 years. Subjects were expected to be in one of the five health conditions – healthy (cancer-free), TNM Stage 1, TNM Stage 2, TNM Stage 3 and untreatable terminal stage (Stage 4). Early or advanced stage patients (stage 1, 2, and 3) received curative treatment with a stage specific cancer recurrence possibility after a mean duration of 2 years [6, 7], while terminal

So JBY, *et al. Gut* 2021; 70:829–837. doi: 10.1136/gutjnl-2020-322065

cancer patients (stage 4) received only palliative care with a conditional life expectancy of 1 year [6]. As prognosis of the cancer recurrence is poor, patients diagnosed with recurrence were assumed untreatable (equivalent to stage 4) and were given palliative care. Only gastric cancer related mortality was compared as the background mortality due to natural or other causes was expected to be similar in both the scenarios.

The current practice of no screening evaluates the costs, health impacts and mortality as per the current diagnosis rate of gastric cancer in this specific population cohort. We used the published age-specific annual incidence rates of gastric cancer and stage of diagnosis among Chinese Males in Singapore. In the current practice of no-screening we did not account for the cost of false positive endoscopies and the diagnostic expenses of only the true cancer cases was considered, which was a conservative assumption favoring no-screening, similar to the assumption in earlier studies [8, 9]. The proposed mass screening program on the other hand was expected to screen the compliant cohort, identify the cancer cases early due to regular screening and computes the cost, health impacts and mortality accordingly. The subjects tested negative in the screening program will include both healthy cases and missed cancer cases. The missed cancer patients were expected to experience the consequences of treatment delays – disease progression, impact on cost and quality-of-life and an increased mortality, as the cancer would progress in them undiagnosed and untreated and the healthy cases are expected to remain healthy with a possibility of developing gastric cancer in future. Cancers missed in the mass screening program were considered to progress to advanced stages and are expected to be diagnosed at stage 4 due to presentation of symptoms in clinics where they are investigated by endoscopy and biopsy. A 1-year progression time was estimated between the consecutive cancer stages, i.e. a missed stage 1 cancer is expected to progress to stage 2 and then to stage 3 and stage 4 with a one year gap each between the successive stages. Stage 4 patients which were missed in diagnosis were expected to be diagnosed after a mean time of 2 months due to worsening of symptoms.
The compliance rate for mass screening was assumed to be 45% as per the reported compliance in national gastric cancer screening programs in Korea [10] with the non-compliant group expected to behave similar to the current strategy of no-screening. As the miRNA test is simpler to administer and potentially cheaper than the currently used screening methods of UGIS, X-ray or endoscopy, it is hoped to improve the population compliance rate. Thus the performance of the mass screening program across a range of compliance rates (45% - 100%) was also evaluated. All costs quoted in US dollars have been calculated based on the exchange rate of $1.38 Singapore dollars to 1 USD as per exchange rates in July 2017. All costs and health benefits were discounted at an annual rate of 3%.

- **Scenario and sensitivity analysis:** The cost advantages and non-invasive nature of miRNA testing may increase patient compliance with screening relative to current technologies. Scenarios that capture a range of improved compliance rates (45%–100%) were modeled to evaluate the possible impact on early diagnosis (**Figure S5**). An extensive sensitivity analyses was conducted by varying the values of key parameters—endoscopy cost, miRNA test costs, miRNA test specificity/sensitivity by cancer stages (stages 1, 2, 3, 4), QoL values by cancer stages (stage 1, 2, 3, 4), cancer recurrence rates by the stage at diagnosis and average annual incidence of gastric cancer—to evaluate model robustness at a Willingness-to-Pay threshold of 50,000 USD/ QALY (**Figure S8, 9, 10**).

- **Treatment Protocols for Cancer Treatment and Related Costs:**
Stage-specific treatment protocols and average medical expenditures for gastric cancer were obtained from the National University Hospital and expert opinions of clinicians based on current

practices in Singapore (Supplementary Table11). Patients diagnosed with gastric cancer undergo staging investigation, which includes Computerized Tomography (CT), Chest X-Ray (CX-R), Endoscopic Ultrasound (EUR) and a specialist consultation (including the cost of nurse counseling and an estimated round-trip transport). Curative treatment administered to stage 1, 2, and 3 cancer patients includes surgery (total/ partial gastrectomy) and hospital stay of 12 days. Stage 3 patients undergo an additional chemo-radiotherapy (5 follow-ups) and radiotherapy sessions (5 sessions/ week for 5 weeks). Palliative care for stage 4 patients includes bypass surgery (30%), endoscopic stenting (6%), palliative chemotherapy-5 sessions (16%) and conservative treatment (2x specialist visits) (48%) with an appropriate hospital stay (12 days in surgery cases and 2.5 days for cases with no surgery). Patients are also expected to adhere to follow up visits (average 2.2 visits/ 5 years) and repeat CT and CXR (average 1.5per year for 5years after the diagnosis).
The miRNA panel test cost in Singapore has been assumed to be USD 30 with an additional 10% for handling and administrative purposes. However, the cost of organizing mass screening has not been included. Costs and QALYs were presented on a present-value basis, with an annual discount rate of 3%. All the diagnosed cases are expected to undergo a biopsy examination. Total costs have been evaluated inclusive of Goods and Services Tax (GST) and without considering any government subsidy.

- **Quality of life values:**
  Stage-specific EQ-5D quality-of-life (QoL) index measures were obtained from a previous local study [11] performed on Chinese gastric cancer patients in National University Hospital, Singapore (**Table S6**). A diagnosed patient is expected to be immediately started on treatment, and experience the diagnosed stage-specific QoL for 1 year with a 6-month additional decrease in QoL due to the initial surgery referred as disutility. After one year of treatment, the patient is expected to enjoy a QoL equivalent to an asymptomatic patient (similar to stage 1 cancer) for the remaining time until faced with any recurrence, which would subsequently drop the QoL to stage 4 equivalent (**Table S6**).

- **Test Characteristics:**
  Test characteristics for diagnostic endoscopy with biopsy for the suspected cases (sensitivity: 93%, specificity: 100%) has been resourced from a study evaluating diagnostic accuracy through a retrospective study among gastric cancer patients [12]. Biopsy is believed to be perfect with 100% sensitivity and specificity. The miRNA stage –specific sensitivity and specificity as estimated from the Singapore Discovery Cohort have been considered as the base-case value.

- **Estimation of population prevalence of undiagnosed gastric cancer:**
  As the study aim to identify the benefits of early diagnosis of gastric cancer, it is essential to calculate the population prevalence of undiagnosed gastric cancer cases in the target group. The current annual age-specific incidence rate is 57 cancers per 100,000 in this population cohort [4] with a stage specific distribution of stage 1: 2:3: 4:: 18%:12%:27%:43% . Based on the assumption of 1 year time for progression of cancer from one stage to another, undiagnosed cancer prevalence in the population cohort (stage 1 and higher) was evaluated individually before every mass screening follow-up. Also, the stage 1 and 2 cancers which currently develop and are diagnosed in between the follow-up years are expected to continue to be diagnosed in both the strategies.

So JBY, *et al. Gut* 2021; 70:829–837. doi: 10.1136/gutjnl-2020-322065

Supplementary Figures

Figure S1. Multi-layered control measures for absolute quantification of miRNA expression

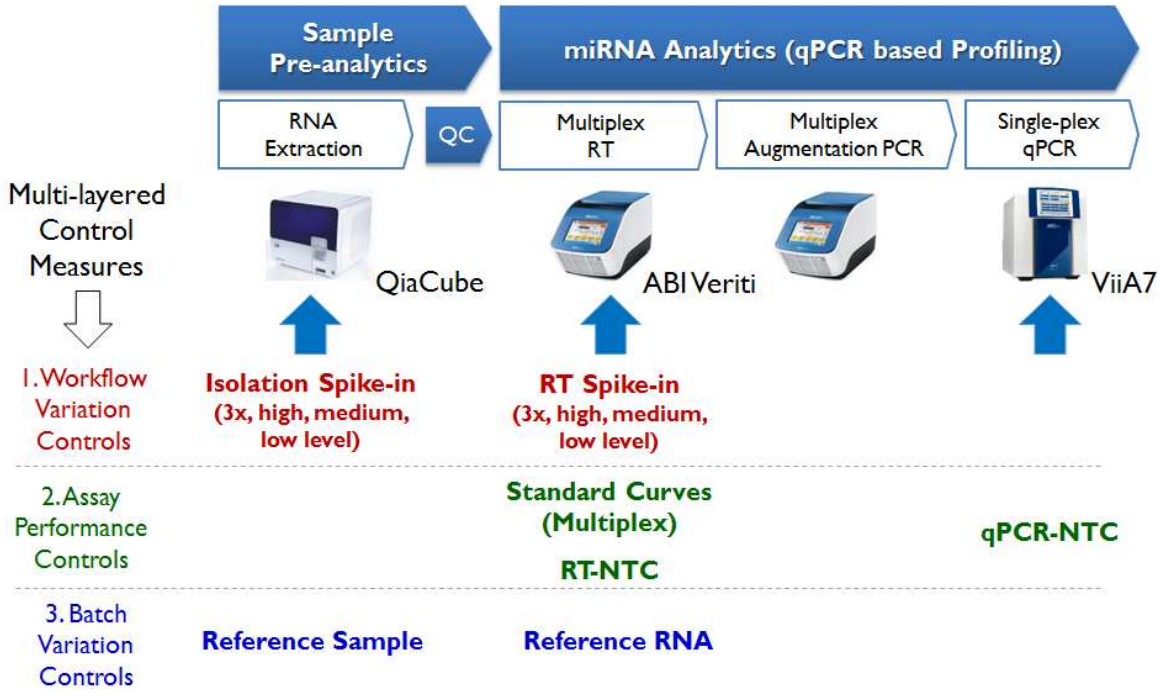


Figure S2. Analytical Validation of MIRXES miRNA RT-qPCR Assays. (A) Analytical Specificity Test. Table below shows the relative detection when assays are challenged with mismatched let-7 family members. **(B) Analytical Reproducibility Test**. Correlation of the expression profiles of 200 miRNAs quantified in 15 control and 15 gastric cancer sera in two independent laboratories (R2> 0.95) using MIRXES miRNA RT-qPCR assays. **(C) Analytical Sensitivity Test.** Comparison with Probe Based Assays. Consistency in the analytical sensitivity of the MIRXES assays was demonstrated by amplification and detection of a 7-log serial dilution of the synthetic templates of 8 miRNAs with low to high AT content. In contrast, probe based assays showed poorer consistency across miRNAs with different AT content.

Figure S2A

	Target	Assay relative detection								
		let-7a	let-7b	let-7c	let-7d	let-7e	let-7f	let-7g	let-7i	miR-98
UGAGGUAGUAGGUUGUAUAGUU	let-7a	100.00%	0.00%	0.10%	1.20%	0.00%	0.50%	0.00%	0.00%	0.00%
UGAGGUAGUAGGUUGUGUGGUU	let-7b	0.00%	100.00%	2.20%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
UGAGGUAGUAGGUUGUAUGGUU	let-7c	0.80%	1.60%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
AGAGGUAGUAGGUUGCAUAGUU	let-7d	0.40%	0.00%	0.10%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
UGAGGUAGGAGGUUGUAUAGUU	let-7e	0.10%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
UGAGGUAGUAGAUUGUAUAGUU	let-7f	0.10%	0.00%	0.00%	0.00%	0.10%	100.00%	0.00%	0.00%	0.00%
UGAGGUAGUAGUUUGUACAGUU	let-7g	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
UGAGGUAGUAGUUUGUGCUGUU	let-7i	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
UGAGGUAGUAAGUUGUAUUGUU	miR-98	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%

Figure S2B

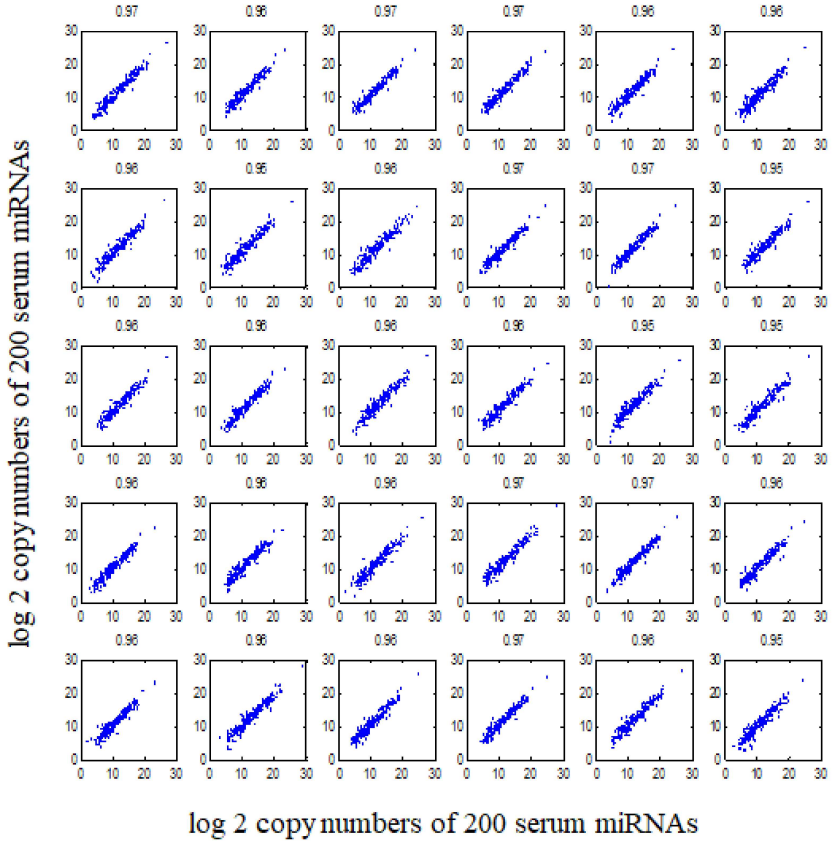
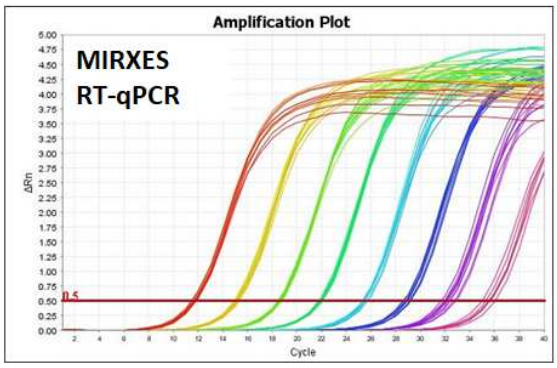


Figure S2C

#	miRNA	Sequence	AT%
1	hsa-miR-99b-5p	CACCCGUAGAACCGACCUUGCG	36.4%
2	hsa-miR-133a-3p	UUUGGUCCCCUUCAACCAGCUG	45.5%
3	hsa-miR-593-3p	UGUCUCUGCUGGGGUUUCU	47.4%
4	hsa-miR-151a-5p	UCGAGGAGCUCACAGUCUAGU	47.6%
5	hsa-miR-20b-5p	CAAAGUGCUCAUAGUGCAGGUAG	52.2%
6	hsa-miR-122-5p	UGGAGUGUGACAAUGGUGUUUG	54.5%
7	hsa-miR-377-3p	AUCACACAAAGGCAACUUUUGU	59.1%
8	hsa-miR-451a	AAACCGUUACCAUUACUGAGUU	63.6%



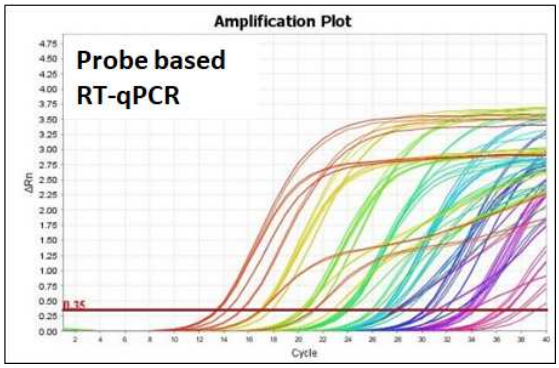


Figure S3. (A) Top up-regulated and down-regulated miRNAs in gastric cancer. The boxplots and ROC curves of top-ranked (based on AUC) up-regulated (miR-142-5p) and down-regulated (miR-99b-5p) miRNAs in all the gastric cancer subjects (regardless of subtypes and stages) compared to normal subjects; the expression levels (copy/ml) were presented in log2 scale. The boxplot presented the 25th, 50th, and 75th percentiles in the distribution of expression levels. AUC: area under the ROC curve. **(B) Top up-regulated and down-regulated miRNAs between normal and various stages of gastric cancers.** The ROC curves of top-ranking (based on AUC) up-regulated (row 1) and down-regulated (row 2) miRNAs in the various stages of gastric cancers subjects compared to the normal subjects. AUC: area under the ROC curve.

Figure S3A

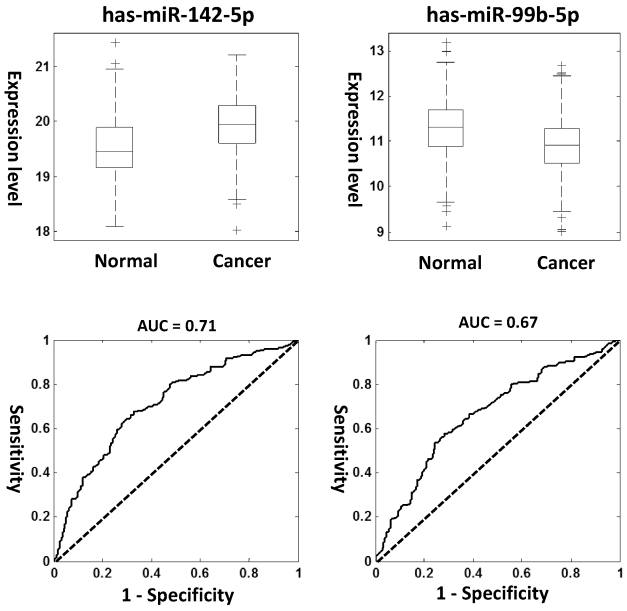


Figure S3B

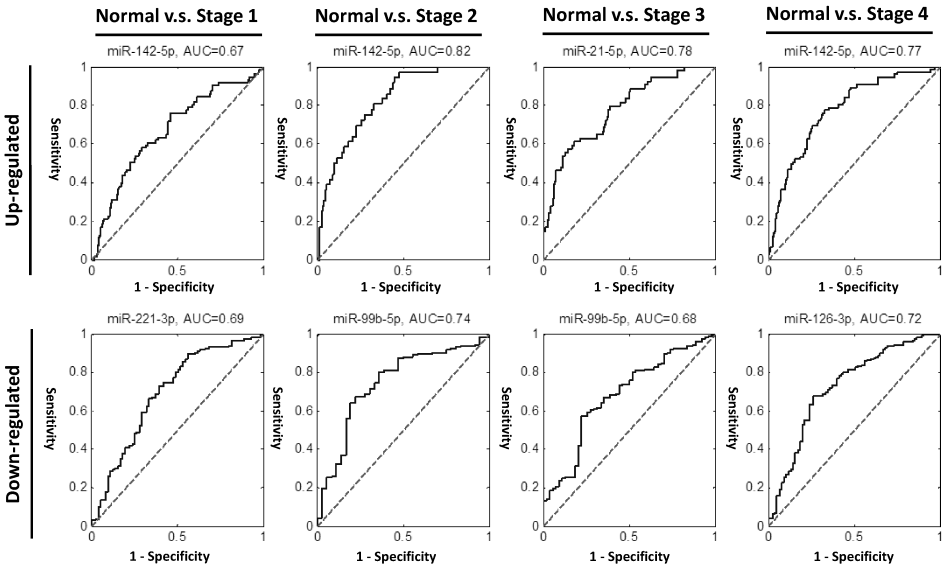

Normal v.s. Stage 1
Normal v.s. Stage 2
Normal v.s. Stage 3
Normal v.s. Stage 4
Up-regulated
Down-regulated
miR-142-5p, AUC=0.67
miR-142-5p, AUC=0.82
miR-21-5p, AUC=0.78
miR-142-5p, AUC=0.77
miR-221-3p, AUC=0.69
miR-99b-5p, AUC=0.74
miR-99b-5p, AUC=0.68
miR-126-3p, AUC=0.72
Sensitivity
1 - Specificity

Figure S4. Differentially Expressed miRNAs between Various Gastric Cancer Subtypes. Boxplot of 7 miRNAs with p-values lower than 0.05 for three different subtypes based on two-way Anova test after false discovery rate (FDR) correction (Bonferroni method). The p-values were shown above the miRNAs. The boxplot presented the 25th, 50th, and 75th percentiles in the distribution of log2 scale expression levels (copy/ml). For each miRNA, the significant levels between various subtypes were calculated based on Bonferroni adjustment to compensate for multiple comparisons. *: p-value < 0.05; **: p-value < 0.01; ***: p-value < 0.001. Five (hsa-miR-27a-3p, hsa-miR-338-5p, hsa-miR-181d, hsa-miR-146b-5p, hsa-miR-30e-3p) of these 7 miRNAs were found to be up-regulated and 2 were found to be down-regulated (hsa-miR-21-3p, hsa-miR-1226-3p) in the diffuse subtype when compared to the intestinal subtype. In addition, the expression level of these 7 miRNAs in the mixed subtype were found to be similar to the diffuse subtype except for hsa-miR-146b-5p, where the expression level in diffuse subtype was higher than the other two subtypes (middle lower panel).

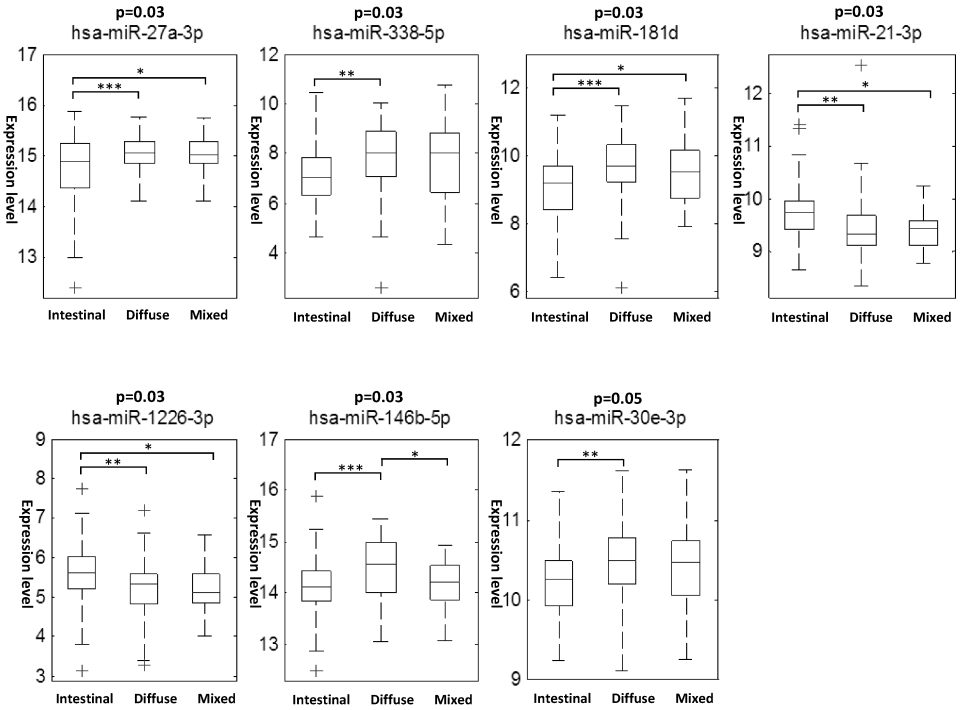


Figure S5. The impact of population compliance in screening program on early cancer diagnosis. With increasing compliance in the mass screening program, more cancer cases are expected to be diagnosed in early stages as compared to late stages. The figure below represents the expected pattern of early diagnosis with higher compliance rates. The ICER which is the cost spent to gain 1 additional QALY in the life of patient is independent of the compliance rate and is equivalent to USD 28,931/ QALY in all scenarios. *Abbreviations used- QALY: Quality-adjusted life years, ICER: Incremental cost-effectiveness ratio, Mn: Million, USD: United States Dollars*

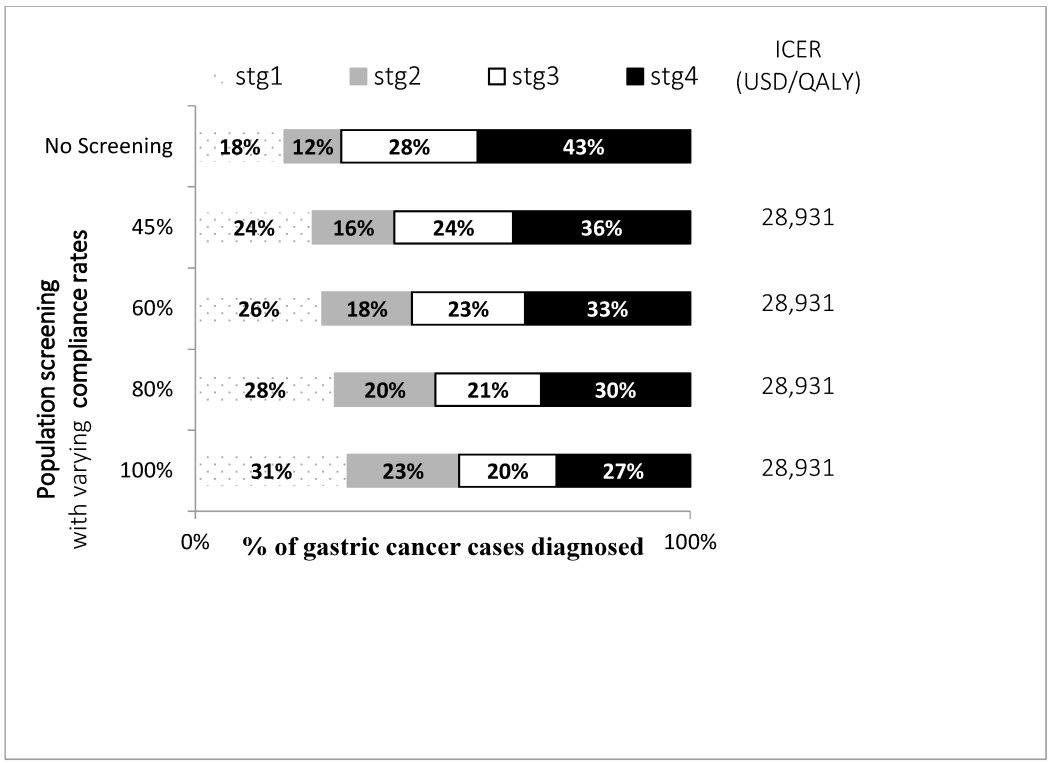


Figure S6. Comparison of Serum miRNA Biomarkers for Gastric, Breast and Ovarian Cancers. Three independent studies have been conducted to investigate serum miRNA expression changes between gastric, breast and ovarian cancers with their corresponding control populations. All measurements were performed using identical assays and workflows. The overlaps of up- and down-regulated serum miRNA biomarkers for these three cancers were presented in the venn diagram below. While there are some overlaps among the 3 cancers, distinct miRNAs changes specific to each cancer were observed. The heatmap below illustrates distinct serum miRNA expression in breast and gastric cancer patients.

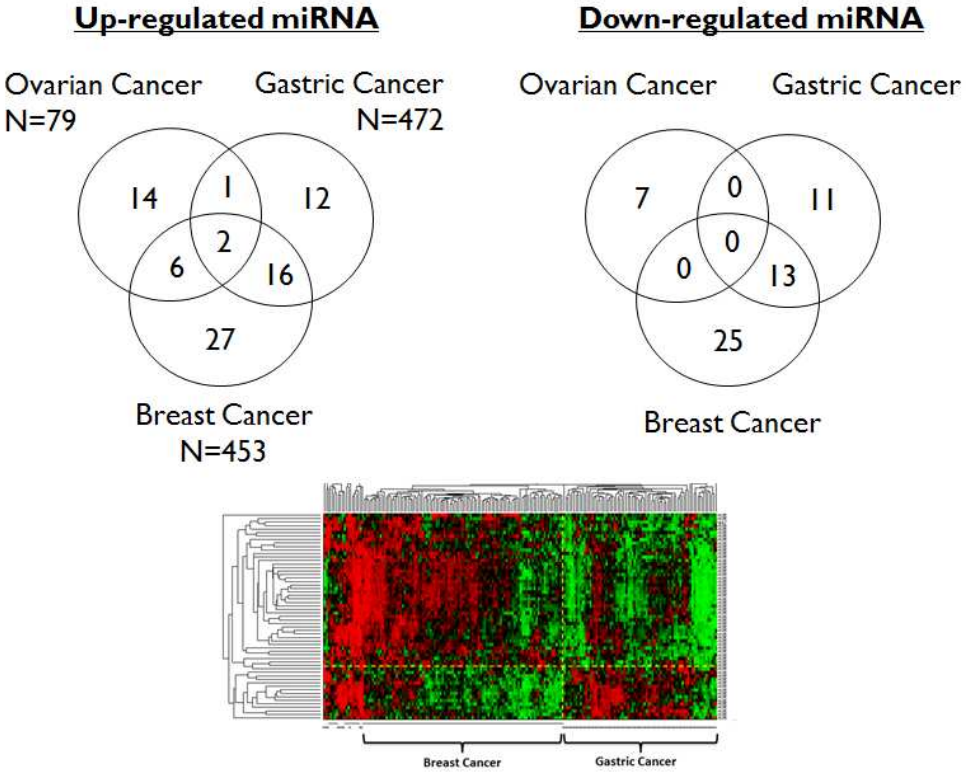


Figure S7. Markov-Decision Tree Model for evaluating the cost-effectiveness. The Markov decision model built has a 25-year time horizon and represents the movement of subjects from one health state to another in both the strategies. The patients who are diagnosed with cancer are treated as per the stage of diagnosis. Cancer recurrence is considered the only reason of treatment failure with all the recurrent cases being fatal. Subject is expected to exit the cohort at age 75. Health and cost parameters corresponding to each state are indicated at every step. The Markov model considers a 3% discount rate for both cost and health benefits and calculates values in their net present value.

Figure S7(A) Decision Tree Model for Strategy 1 – No screening. In current practice the subjects are diagnosed in clinic and treated as per the stage of diagnosis. The subjects in the population cohort would continue living healthy until being diagnosed as per the annual incidence rate reported in the Singapore cancer registry.

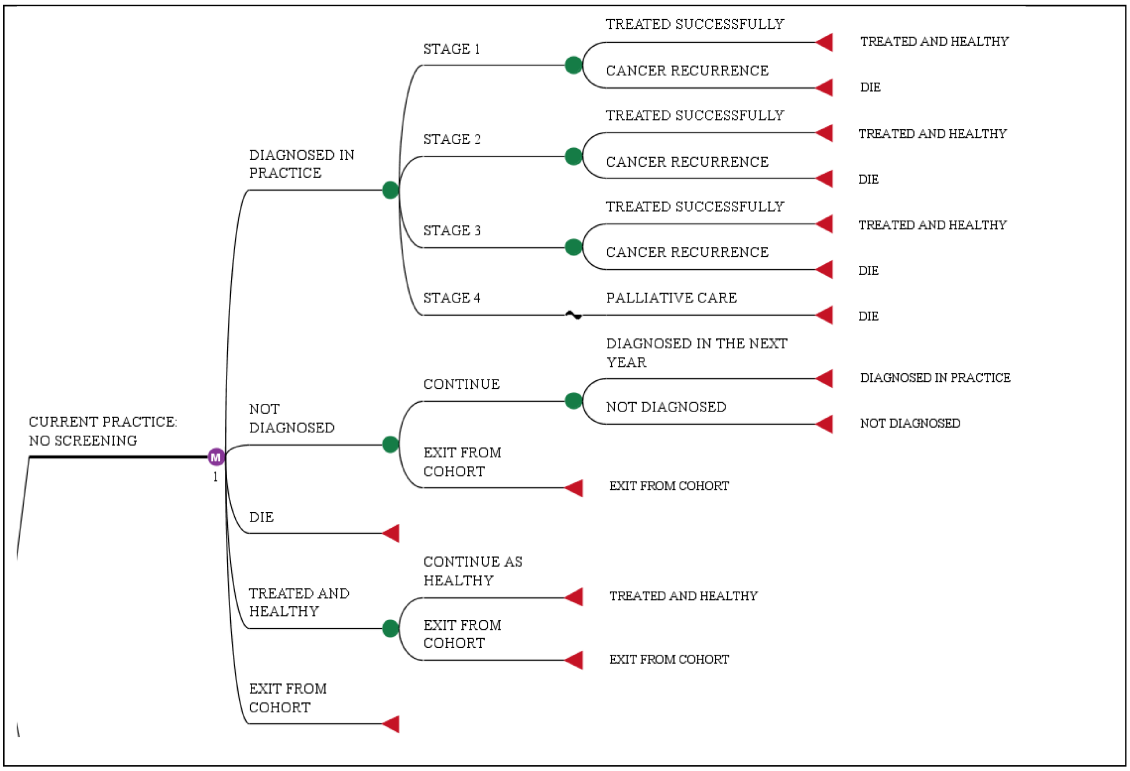


Figure S7(B) Decision Tree Model for Strategy2: mass screening with miRNA followed by endoscopy for diagnosis confirmation for test positive subjects and a 3-year follow up for test negative subjects. The subjects compliant to the mass screening program are made to undergo a miRNA-based blood test. If the miRNA test is positive, the subjects will undergo endoscopy and biopsy to confirm the cancer and if the miRNA test is negative, subjects are considered healthy until new cancers are diagnosed over time. Among this healthy group, there would be cancer cases missed by miRNA test. The missed cancer cases would progress to advanced stages and are expected to be diagnosed with symptoms at stage 4 in the clinic. The remaining healthy cases group is followed up with miRNA test every 3 years.

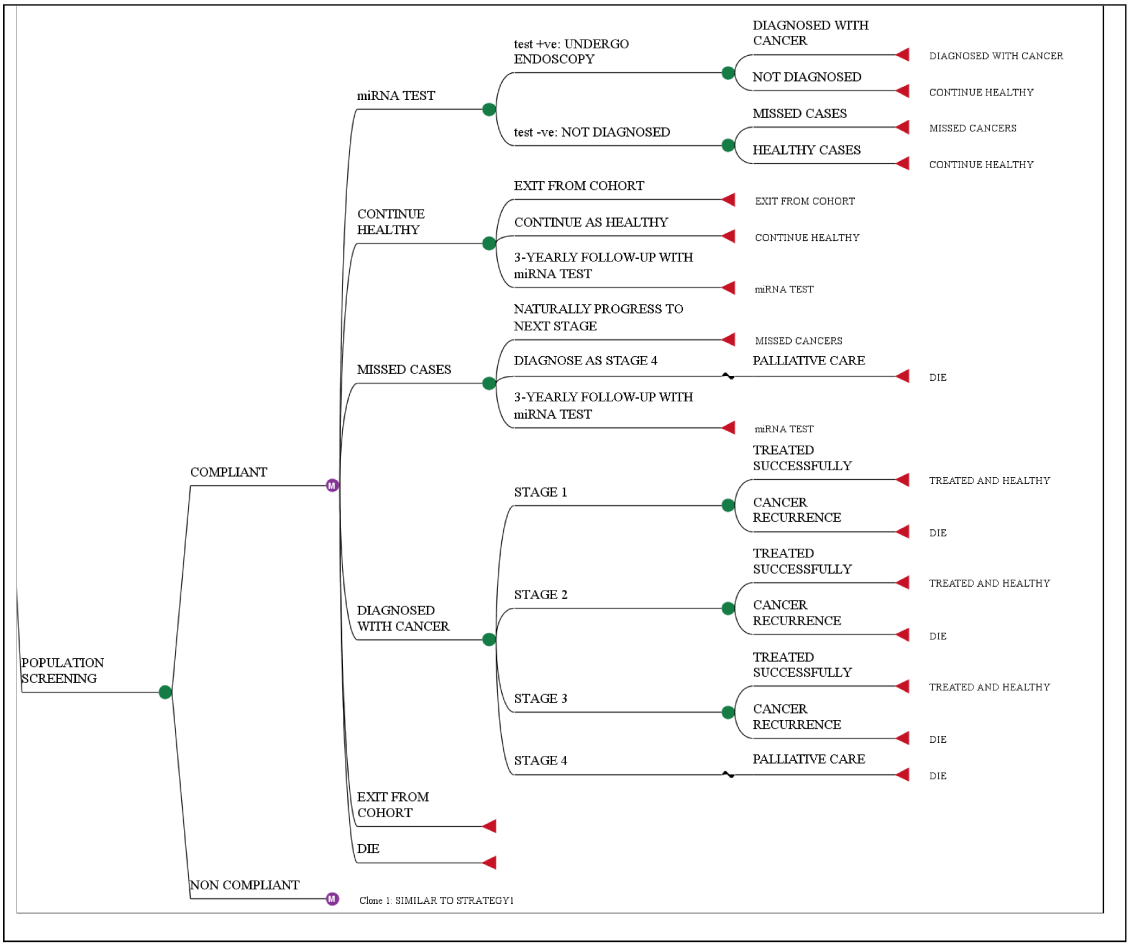


Figure S8. Sensitivity analysis for mass screening of Singaporean Chinese Males (50-75 years). We have performed one-way sensitivity analysis of many key variables to identify the impact of variable uncertainty on the Incremental Cost Effectiveness Ratio (ICER). The figure below shows all variables with their sensitivity ranges on Y axis and the ICER values on X axis. The range of values that were examined is shown in parentheses, with the value giving the lower ICER listed first. The graph represents the possible variation in ICER due to variable uncertainty, with the most significant variables at top. The solid vertical line indicates the ICER of 28,931 USD/QALY for the base-case scenario while the dash line indicates the threshold of ICER 50,000 USD/QALY. Three significant variables were identified which are: miRNA test cost, specificity of miRNA test and sensitivity of miRNA test for stage 1 patients. *Abbreviations used- QALY: Quality adjusted life years, ICER: Incremental cost-effectiveness ratio*

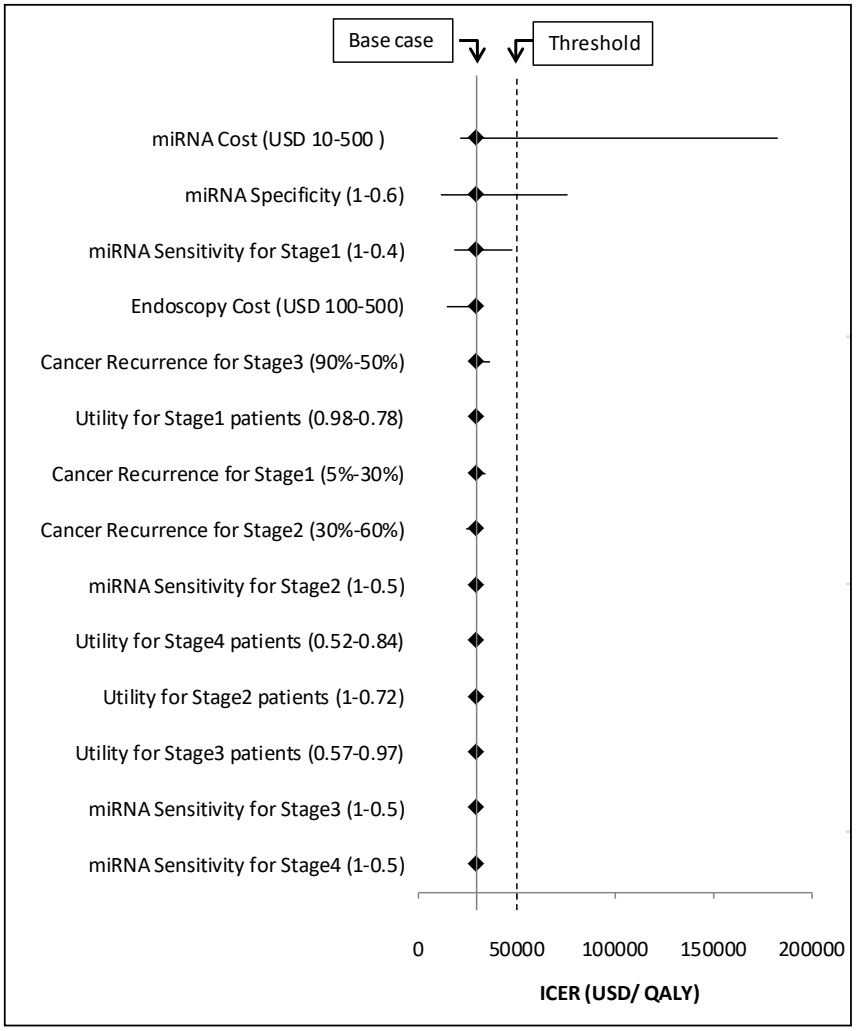


Figure S9. One-way sensitivity analysis for variables found significant for cost-effectiveness evaluation. The variation in ICER within the range of values of the three significant variables has been shown below individually. Also, their threshold value, i.e. value at which the ICER is equivalent to USD 50,000/QALY has been highlighted. The strategy would be cost-effective only with an ICER < 50,000 USD/ QALY. For each sensitivity analysis below, it is assumed that the rest of variables remain constant as described in base-case scenario.

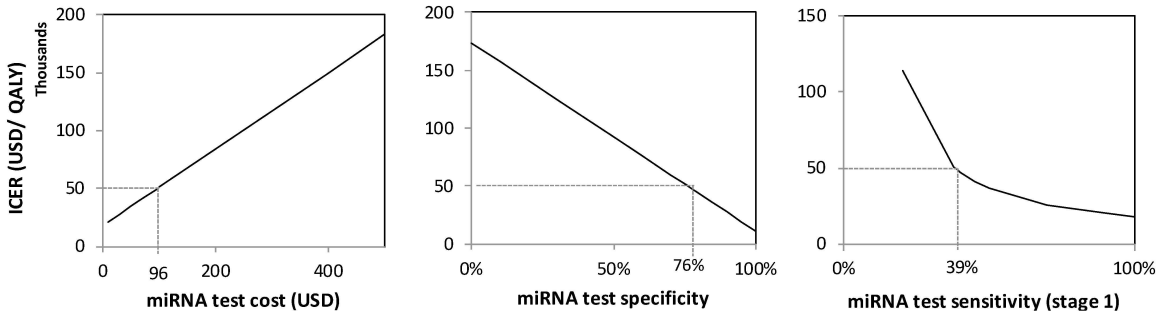


Figure S10. Sensitivity analysis of the cost-effectiveness of screening program with varying gastric cancer incidence. This sensitivity analysis has been performed to identify the cancer incidence which makes screening programs in the population cost-effective. The graph below reports the ICER at the different incidence rates ranging from 0.01% - 0.5%. The incidence rate reported is the average annual incidence for the target population. The analysis found the screening program to be cost-effective at an average incidence higher than 0.05% based on the cost-effectiveness threshold of US$50,000/QALY. The target population of Singaporean Chinese males (50-75 years) is estimated to have an average annual gastric cancer incidence of 0.057% at the beginning of the analysisas per the 2016 population cohort statistics. With ageing, the average incidence rate of the cohort has been considered to increase.

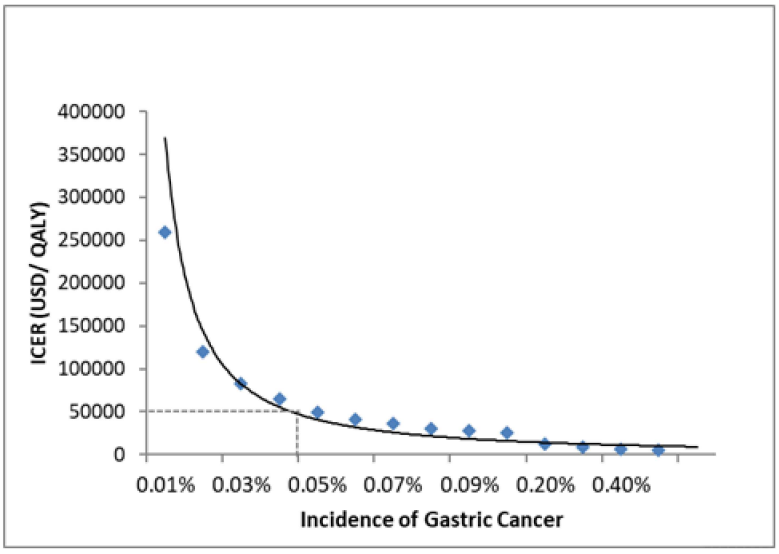


Table S1. Identity and Sequence of 191 Reliable Detected Mature miRNA. 191 mature miRNA were reliable detected in the serum samples. The definition of "reliably detected" was that at least 90% of the serum samples had a concentration higher than 500 copies per ml. The miRNAs were named according to the miRBase V18 release.

Identity	Sequence
hsa-miR-99b-5p	CACCCGUAGAACCGACCUUGCG
hsa-miR-486-5p	UCCUGUACUGAGCUGCCCCGAG
hsa-miR-23b-3p	AUCACAUUGCCAGGGAUUACC
hsa-miR-140-3p	UACCACAGGGUAGAACCACGG
hsa-miR-101-3p	UACAGUACUGUGAUAACUGAA
hsa-miR-107	AGCAGCAUUGUACAGGGCUAUCA
hsa-miR-130b-3p	CAGUGCAAUGAUGAAAGGGCAU
hsa-miR-369-3p	AAUAAUACAUGGUUGAUCUUU
hsa-miR-133a	UUUGGUCCCCUUCAACCAGCUG
hsa-miR-222-3p	AGCUACAUCUGGCUACUGGGU
hsa-miR-320d	AAAAGCUGGGUUGAGAGGA
hsa-miR-30a-5p	UGUAAACAUCCUCGACUGGAAG
hsa-miR-181a-5p	AACAUUCAACGCUGUCGGUGAGU
hsa-miR-140-5p	CAGUGGUUUUACCCUAUGGUAG
hsa-miR-425-3p	AUCGGGAAUGUCGUGUCCGCCC
hsa-miR-106b-3p	CCGCACUGUGGGUACUUGCUGC
hsa-miR-192-5p	CUGACCUAUGAAUUGACAGCC
hsa-miR-10a-3p	CAAAUUCGUAUCUAGGGGAAUA
hsa-miR-17-5p	CAAAGUGCUUACAGUGCAGGUAG
hsa-miR-590-5p	GAGCUUAUUCAUAAAAGUGCAG
hsa-miR-1299	UUCUGGAAUUCUGUGUGAGGGA
hsa-miR-365a-3p	UAAUGCCCCUAAAAAUCCUUAU
hsa-miR-500a-5p	UAAUCCUUGCUACCUGGGUGAGA
hsa-miR-32-5p	UAUUGCACAUUACUAAGUUGCA
hsa-miR-340-5p	UUAUAAAGCAAUGAGACUGAUU
hsa-miR-374b-5p	AUAUAAUACAACCUGCUAAGUG
hsa-miR-27a-3p	UUCACAGUGGCUAAGUUCCGC
hsa-miR-627	GUGAGUCUCUAAGAAAAGAGGA
hsa-miR-539-5p	GGAGAAAUUAUCCUUGGUGUGU
hsa-miR-342-5p	AGGGGUGCUAUCUGUGAUUGA
hsa-miR-484	UCAGGCUCAGUCCCCUCCCGAU
hsa-miR-132-3p	UAACAGUCUACAGCCAUGGUCG
hsa-miR-379-5p	UGGUAGACUAUGGAACGUAGG
hsa-miR-125a-3p	ACAGGUGAGGUUCUUGGGAGCC
hsa-miR-29a-3p	UAGCACCAUCUGAAAUCGGUUA

hsa-miR-363-3p	AAUUGCACGGUAUCCAUCUGUA
hsa-miR-376b	AUCAUAGAGGAAAAUCCAUGUU
hsa-miR-589-5p	UGAGAACCACGUCUGCUCUGAG
hsa-miR-432-5p	UCUUGGAGUAGGUCAUUGGGUGG
hsa-miR-1280	UCCCACCGCUGCCACCC
hsa-miR-103a-3p	AGCAGCAUUGUACAGGGCUAUGA
hsa-miR-122-5p	UGGAGUGUGACAAUGGUGUUUG
hsa-miR-93-5p	CAAAGUGCUGUUCGUGCAGGUAG
hsa-miR-25-3p	CAUUGCACUUGUCUCGGUCUGA
hsa-miR-9-5p	UCUUUGGUUAUCUAGCUGUAUGA
hsa-miR-579	UUCAUUUGGUAUAAACCGCGAUU
hsa-miR-136-3p	CAUCAUCGUCUCAAAUGAGUCU
hsa-miR-146a-5p	UGAGAACUGAAUUCCAUGGGUU
hsa-miR-144-5p	GGAUAUCAUCAUAUACUGUAAG
hsa-miR-15a-5p	UAGCAGCACAUAAUGGUUUGUG
hsa-miR-150-5p	UCUCCCAACCCUUGUACCAGUG
hsa-miR-152	UCAGUGCAUGACAGAACUUGG
hsa-miR-29c-5p	UGACCGAUUUCUCCUGGUGUUC
hsa-miR-320c	AAAAGCUGGGUUGAGAGGGU
hsa-miR-127-3p	UCGGAUCCGUCUGAGCUUGGCU
hsa-miR-331-5p	CUAGGUAUGGUCCCAGGGAUCC
hsa-miR-378a-3p	ACUGGACUUGGAGUCAGAAGG
hsa-miR-374a-5p	UUAUAAUACAACCUGAUAAGUG
hsa-miR-409-3p	GAAUGUUGCUCGGUGAACCCCU
hsa-miR-411-3p	UAUGUAACACGGUCCACUAACC
hsa-miR-505-3p	CGUCAACACUUGCUGGUUUCCU
hsa-miR-628-5p	AUGCUGACAUAUUUACUAGAGG
hsa-miR-629-3p	GUUCUCCCAACGUAAGCCCAGC
hsa-miR-4732-3p	GCCCUGACCUGUCCUGUUCUG
hsa-miR-501-5p	AAUCCUUUGUCCCUGGGUGAGA
hsa-miR-616-5p	ACUCAAAACCCUUCAGUGACUU
hsa-miR-454-3p	UAGUGCAAUAUUGCUUAUAGGGU
hsa-miR-485-3p	GUCAUACACGGCUCUCCUCUCU
hsa-miR-133b	UUUGGUCCCCUUCAACCAGCUA
hsa-miR-186-5p	CAAAGAAUUCUCCUUUUGGGCU
hsa-miR-20b-5p	CAAAGUGCUCAUAGUGCAGGUAG
hsa-miR-30d-5p	UGUAAACAUCCCCGACUGGAAG
hsa-miR-375	UUUGUUCGUUCGGCUCGCGUGA
hsa-miR-16-5p	UAGCAGCACGUAAAUAUUGGCG
hsa-miR-106b-5p	UAAAGUGCUGACAGUGCAGAU
hsa-miR-139-5p	UCUACAGUGCACGUGUCUCCAG

hsa-miR-141-3p	UAACACUGUCUGGUAAAGAUGG
hsa-miR-185-5p	UGGAGAGAAAGGCAGUUCCUGA
hsa-miR-181b-5p	AACAUUCAUUGCUGUCGGUGGGU
hsa-miR-199a-3p	ACAGUAGUCUGCACAUUGGUUA
hsa-miR-19b-3p	UGUGCAAAUCCAUGCAAAACUGA
hsa-miR-148b-3p	UCAGUGCAUCACAGAACUUUGU
hsa-miR-29b-3p	UAGCACCAUUUGAAAUCAGUGUU
hsa-miR-338-5p	AACAAUAUCCUGGUGCUGAGUG
hsa-miR-584-5p	UUAUGGUUUGCCUGGGACUGAG
hsa-miR-382-5p	GAAGUUGUUCGUGGUGGAUUCG
hsa-miR-151a-3p	CUAGACUGAAGCUCCUUGAGG
hsa-miR-1290	UGGAUUUUUGGAUCAGGGA
hsa-miR-200b-3p	UAAUACUGCCUGGUAAUGAUGA
hsa-miR-411-5p	UAGUAGACCGUAUAGCGUACG
hsa-miR-126-5p	CAUUAUUACUUUUGGUACGCG
hsa-miR-101-5p	CAGUUAUCACAGUGCUGAUGCU
hsa-miR-125b-5p	UCCCUGAGACCCUAACUUGUGA
hsa-miR-362-5p	AAUCCUUGGAACCUAGGUGUGAGU
hsa-miR-197-3p	UUCACCACCUUCUCCACCCAGC
hsa-miR-221-3p	AGCUACAUUGUCUGCUGGGUUUC
hsa-miR-501-3p	AAUGCACCCGGGCAAGGAUUCU
hsa-miR-671-3p	UCCGGUUCUCAGGGCUCCACC
hsa-miR-181a-2-3p	ACCACUGACCGUUGACUGUACC
hsa-miR-9-3p	AUAAAGCUAGAUAACCGAAAGU
hsa-miR-452-5p	AACUGUUUGCAGAGGAAACUGA
hsa-miR-598	UACGUCAUCGUUGUCAUCGUCA
hsa-miR-320b	AAAAGCUGGGUUGAGAGGGCAA
hsa-miR-328	CUGGCCCUCUCUGCCCUUCCGU
hsa-miR-650	AGGAGGCAGCGCUCUCAGGAC
hsa-miR-134	UGUGACUGGUUGACCAGAGGGG
hsa-miR-130a-3p	CAGUGCAAUGUUAAAAGGGCAU
hsa-miR-21-5p	UAGCUUAUCAGACUGAUGUUGA
hsa-miR-424-5p	CAGCAGCAAUUCAUGUUUUGAA
hsa-miR-99a-5p	AACCCGUAGAUCCGAUCUUGUG
hsa-miR-18a-3p	ACUGCCCUAAGUGCUCCUUCUGG
hsa-miR-195-5p	UAGCAGCACAGAAAUAUUGGC
hsa-miR-205-5p	UCCUUCAUUCCACCGGAGUCUG
hsa-miR-206	UGGAAUGUAAGGAAGUGUGUGG
hsa-miR-500a-3p	AUGCACCUGGGCAAGGAUUCUG
hsa-miR-18b-5p	UAAGGUGCAUCUAGUGCAGUUAG
hsa-miR-181d	AACAUUCAUUGUUGUCGGUGGGU

hsa-miR-339-3p	UGAGCGCCUCGACGACAGAGCCG
hsa-miR-93-3p	ACUGCUGAGCUAGCACUUCCCG
hsa-miR-10b-5p	UACCCUGUAGAACCGAAUUUGUG
hsa-miR-497-5p	CAGCAGCACACUGUGGUUUGU
hsa-miR-27b-3p	UUCACAGUGGCUAAGUUCUGC
hsa-miR-128	UCACAGUGAACCGGUCUCUUU
hsa-miR-183-5p	UAUGGCACUGGUAGAAUUCACU
hsa-miR-22-3p	AAGCUGCCAGUUGAAGAACUGU
hsa-miR-26a-5p	UUCAAGUAAUCCAGGAUAGGCU
hsa-miR-223-3p	UGUCAGUUUGUCAAAUACCCCA
hsa-miR-629-5p	UGGGUUUACGUUGGGAGAACU
hsa-miR-92a-3p	UAUUGCACUUGUCCCGGCCUGU
hsa-miR-29b-2-5p	CUGGUUUCACAUGGUGGCUUAG
hsa-miR-21-3p	CAACACCAGUCGAUGGGCUGU
hsa-miR-199a-5p	CCCAGUGUUCAGACUACCUGUUC
hsa-miR-148a-3p	UCAGUGCACUACAGAACUUUGU
hsa-miR-193a-5p	UGGGUCUUUGCGGGCGAGAUGA
hsa-miR-27a-5p	AGGGCUUAGCUGCUUGUGAGCA
hsa-miR-200c-3p	UAAUACUGCCGGGUAAUGAUGGA
hsa-miR-20a-5p	UAAAGUGCUUAUAGUGCAGGUAG
hsa-miR-194-5p	UGUAACAGCAACUCCAUGUGGA
hsa-miR-532-3p	CCUCCCACACCCAAGGCUUGCA
hsa-miR-19a-3p	UGUGCAAAUCUAUGCAAAACUGA
hsa-miR-142-5p	CAUAAAGUAGAAAGCACUACU
hsa-miR-144-3p	UACAGUAUAGAUGAUGUACU
hsa-miR-145-5p	GUCCAGUUUUCCCAGGAAUCCCU
hsa-miR-10a-5p	UACCCUGUAGAUCCGAAUUUGUG
hsa-miR-23a-3p	AUCACAUUGCCAGGGAUUUCC
hsa-miR-23a-5p	GGGGUUCCUGGGGAUGGGAUUU
hsa-miR-15b-3p	CGAAUCAUUAUUUGCUGCUCUA
hsa-miR-301a-3p	CAGUGCAAUAGUAUUGUCAAAGC
hsa-miR-660-5p	UACCCAUUGCAUAUCGGAGUUG
hsa-miR-30b-5p	UGUAAACAUCCUACACUCAGCU
hsa-miR-30e-5p	UGUAAACAUCCUUGACUGGAAG
hsa-miR-550a-5p	AGUGCCUGAGGGAGUAAGAGCCC
hsa-miR-425-5p	AAUGACACGAUCACUCCCGUUGA
hsa-miR-4306	UGGAGAGAAAGGCAGUA
hsa-miR-532-5p	CAUGCCUUGAGUGUAGGACCGU
hsa-miR-335-5p	UCAAGAGCAAUAACGAAAAAUGU
hsa-miR-483-5p	AAGACGGGAGGAAAGAAGGGAG
hsa-miR-1226-3p	UCACCAGCCCUGUGUUCCCUAG

hsa-miR-431-5p	UGUCUUGCAGGCCGUCAUGCA
hsa-miR-324-5p	CGCAUCCCCUAGGGCAUUGGUGU
hsa-miR-487b	AAUCGUACAGGGUCAUCCACUU
hsa-miR-451a	AAACCGUUACCAUUACUGAGUU
hsa-miR-493-5p	UUGUACAUGGUAGGCUUUCAUU
hsa-miR-136-5p	ACUCCAUUUGUUUUGAUGAUGGA
hsa-miR-23c	AUCACAUUGCCAGUGAUUACCC
hsa-miR-95	UUCAACGGGUAUUUAUUGAGCA
hsa-miR-423-5p	UGAGGGGCAGAGAGCGAGACUUU
hsa-miR-320e	AAAGCUGGGUUGAGAAGG
hsa-miR-224-5p	CAAGUCACUAGUGGUUCCGUU
hsa-miR-28-3p	CACUAGAUUGUGAGCUCCUGGA
hsa-miR-29c-3p	UAGCACCAUUUGAAAUCGGUUA
hsa-miR-326	CCUCUGGGCCCUUCCUCCAG
hsa-miR-596	AAGCCUGCCCGGCUCCUCGGG
hsa-miR-885-5p	UCCAUUACACUACCCUGCCUCU
hsa-miR-146b-5p	UGAGAACUGAAUUCCAUAGGCU
hsa-miR-34a-5p	UGGCAGUGUCUUAGCUGGUUGU
hsa-miR-330-3p	GCAAAGCACACGGCCUGCAGAGA
hsa-miR-154-5p	UAGGUUAUCCGUGUUGCCUUCG
hsa-miR-191-5p	CAACGGAAUCCCAAAAGCAGCUG
hsa-miR-193b-3p	AACUGGCCCUCAAAGUCCCGCU
hsa-miR-301b	CAGUGCAAUGAUAUUGUCAAAGC
hsa-miR-30e-3p	CUUUCAGUCGGAUGUUUACAGC
hsa-miR-320a	AAAAGCUGGGUUGAGAGGGCGA
hsa-miR-199b-3p	ACAGUAGUCUGCACAUUGGUUA
hsa-miR-502-3p	AAUGCACCUGGGCAAGGAUUCA
hsa-miR-450a-5p	UUUUGCGAUGUGUUCCUAAUAU
hsa-miR-495	AAACAAACAUGGUGCACUUCUU
hsa-miR-126-3p	UCGUACCGUGAGUAAUAAUGCG
hsa-miR-15b-5p	UAGCAGCACAUCAUGGUUUACA
hsa-miR-339-5p	UCCCUGUCCUCCAGGAGCUCACG
hsa-miR-337-5p	GAACGGCUUCAUACAGGAGUU

Table S2. MiRNAs Differentially Expressed between Normal and Gastric Cancer. For the comparison between normal and all gastric cancer subjects (regardless of subtypes and stages), 75 miRNA had p-value lower than 0.01 after FDR correction (Bonferroni method). AUC – area under the receiver operating characteristic curve; fold change – the mean expression level (copy/ml) of miRNA in the cancer population divided by that in the normal population.

Up-regulated miRNAs

miRNA name	AUC	P-value	P-value, FDR correction	Fold change	Novel Observation
miR-101-3p	0.61	1.80E-05	9.50E-05	1.27	Novel
miR-106b-3p	0.66	3.70E-09	4.10E-08	1.13	Novel
miR-106b-5p	0.61	5.10E-04	1.70E-03	1.21	
miR-128	0.62	1.40E-06	8.80E-06	1.16	Novel
miR-1280	0.66	3.10E-09	3.90E-08	1.38	Novel
miR-140-3p	0.62	6.40E-06	3.50E-05	1.2	Novel
miR-140-5p	0.67	6.20E-10	1.20E-08	1.24	Novel
miR-142-5p	0.71	1.90E-14	3.70E-12	1.31	Novel
miR-148a-3p	0.67	2.20E-10	4.80E-09	1.32	Novel
miR-15b-3p	0.62	6.20E-06	3.50E-05	1.32	Novel
miR-17-5p	0.63	1.00E-05	5.50E-05	1.24	
miR-183-5p	0.64	8.80E-07	6.20E-06	1.53	Novel
miR-186-5p	0.59	1.40E-03	3.70E-03	1.11	Novel
miR-18b-5p	0.64	1.50E-07	1.20E-06	1.38	Novel
miR-197-3p	0.68	8.10E-13	5.10E-11	1.32	Novel
miR-19a-3p	0.63	4.90E-07	3.60E-06	1.29	Novel
miR-19b-3p	0.59	1.10E-03	3.00E-03	1.18	Novel
miR-20a-5p	0.65	1.20E-07	1.10E-06	1.35	
miR-20b-5p	0.60	2.90E-04	9.90E-04	1.3	Novel
miR-21-3p	0.60	7.90E-05	3.20E-04	1.13	Novel
miR-21-5p	0.63	2.60E-08	2.80E-07	1.23	
miR-223-3p	0.66	7.00E-10	1.20E-08	1.36	
miR-23a-5p	0.64	1.00E-07	9.20E-07	1.31	Novel
miR-25-3p	0.62	3.40E-05	1.60E-04	1.26	Novel
miR-27a-5p	0.69	1.00E-13	1.00E-11	1.76	
miR-29a-3p	0.61	6.00E-05	2.60E-04	1.17	Novel
miR-29b-2-5p	0.59	4.70E-05	2.10E-04	1.16	Novel
miR-29b-3p	0.61	7.20E-05	3.00E-04	1.18	Novel
miR-29c-3p	0.65	2.00E-09	2.90E-08	1.23	Novel
miR-29c-5p	0.63	1.40E-06	8.80E-06	1.15	Novel
miR-338-5p	0.57	3.70E-03	9.40E-03	1.29	Novel
miR-423-5p	0.60	7.20E-05	3.00E-04	1.18	
miR-424-5p	0.68	7.00E-11	1.90E-09	1.41	Novel

miR-425-3p	0.57	2.20E-03	5.70E-03	1.05	Novel
miR-4306	0.63	1.20E-06	8.00E-06	1.35	Novel
miR-450a-5p	0.67	2.10E-10	4.80E-09	1.53	Novel
miR-486-5p	0.61	9.60E-05	3.70E-04	1.32	Novel
miR-500a-3p	0.60	1.10E-04	4.20E-04	1.2	Novel
miR-501-5p	0.60	9.60E-04	2.80E-03	1.24	Novel
miR-532-3p	0.60	1.90E-04	7.00E-04	1.15	Novel
miR-550a-5p	0.63	9.00E-07	6.20E-06	1.38	Novel
miR-579	0.62	2.20E-05	1.10E-04	1.3	Novel
miR-589-5p	0.63	1.70E-06	1.00E-05	1.18	Novel
miR-590-5p	0.69	3.00E-12	1.40E-10	1.23	Novel
miR-598	0.67	7.10E-12	2.70E-10	1.27	Novel
miR-616-5p	0.65	3.40E-09	4.10E-08	1.35	Novel
miR-627	0.58	7.30E-04	2.30E-03	1.19	Novel
miR-629-3p	0.67	6.10E-11	1.90E-09	1.38	Novel
miR-629-5p	0.63	1.40E-04	5.10E-04	1.5	Novel
miR-93-3p	0.62	5.10E-06	3.00E-05	1.22	Novel
miR-93-5p	0.60	2.30E-04	8.00E-04	1.21	Novel

Down-regulated miRNAs

miRNA name	AUC	P-value	P-value, FDR correction	Fold change	
miR-107	0.65	4.40E-08	4.40E-07	0.8	Novel
miR-122-5p	0.61	8.10E-05	3.20E-04	0.66	Novel
miR-126-3p	0.66	1.70E-09	2.70E-08	0.87	Novel
miR-136-5p	0.61	2.30E-05	1.10E-04	0.72	Novel
miR-139-5p	0.60	8.60E-05	3.40E-04	0.84	Novel
miR-146a-5p	0.59	2.10E-03	5.60E-03	0.89	Novel
miR-154-5p	0.59	8.60E-04	2.60E-03	0.8	Novel
miR-181a-5p	0.60	2.30E-04	8.00E-04	0.92	Novel
miR-193b-3p	0.58	1.20E-03	3.20E-03	0.77	Novel
miR-23c	0.59	8.00E-04	2.40E-03	0.84	Novel
miR-26a-5p	0.60	4.40E-05	2.00E-04	0.86	Novel
miR-30a-5p	0.64	6.70E-08	6.40E-07	0.76	Novel
miR-30b-5p	0.59	9.50E-04	2.80E-03	0.9	Novel
miR-337-5p	0.63	4.80E-07	3.60E-06	0.74	Novel
miR-339-5p	0.64	4.90E-07	3.60E-06	0.79	Novel
miR-382-5p	0.59	1.00E-03	2.90E-03	0.81	Novel

miR-409-3p	0.59	5.00E-04	1.60E-03	0.77	Novel
miR-411-5p	0.6	7.30E-04	2.30E-03	0.74	Novel
miR-485-3p	0.6	6.40E-04	2.00E-03	0.77	Novel
miR-487b	0.59	1.10E-03	3.00E-03	0.76	Novel
miR-495	0.6	2.10E-04	7.40E-04	0.77	Novel
miR-885-5p	0.62	1.90E-05	9.60E-05	0.69	Novel
miR-99a-5p	0.58	2.90E-03	7.50E-03	0.82	Novel
miR-99b-5p	0.67	2.60E-09	3.50E-08	0.78	Novel

Table S3. Summary of Serum / Plasma miRNA Biomarker Studies for Gastric Cancer. The studies that measured the cell-free serum/plasma miRNAs were included in the table. Only the results validated with RT-qPCR were shown. GC: gastric cancer subjects. C: control subjects.

Paper	Up regulated	Down regulated	Method	Samples
Chen Li et al [13]	miR-199a-3p	-	RT-qPCR	Plasma/80GC/70C
Aysegul Gorur et al [14]	-	miR-195-5p	RT-qPCR	Serum/20GC/190C
Hui Cai et al [15]	miR-106b, miR-20a, miR-221	-	RT-qPCR	Plasma/90GC/90C
Mei-Hua Cui et al [16]	miR-181c	-	RT-qPCR	Plasma/30GC/60C
Chen Li et al [17]	miR-199a-3p, miR-151-5p	-	RT-qPCR	Plasma/180GC/100C
Ming-yang Song et al [18]	miR-221, miR-744, miR-376c, miR-191, miR-27a, let-7e, miR-27b, and miR-222	-	RT-qPCR	Serum/82GC/82C
Bo-sheng Li et al [19]	miR-223, miR-21	miR- 218	RT-qPCR	Plasma/60GC/60C
Manuel Valladares-Ayerbes et al [20]	miR-200c	-	RT-qPCR	whole blood/52GC/15C
Wen-Hui Zhang et al [21]	-	miR-375	RT-qPCR	Serum
S. S. Lo et al [22]	miR-370	-	RT-qPCR	Plasma/33GC/33C
M. Tsujiura et al [23]	miR-17-5p, miR-21, miR-106a, miR-106b	let-7a	RT-qPCR	Plasma/69GC/30C
Rui Liu et al [24]	miR-1, miR-20a, miR-27a, miR-34a, miR-423-5p	-	RT-qPCR	Serum/142GC/105C
Hanshao Liu et al [25]	miR-187*,miR-371-5p, miR-378	-	RT-qPCR	Serum/40GC/41C

Table S4. MiRNAs Differentially Expressed between Different Stages of Gastric Cancer. A total of 36 miRNAs with p-value lower than 0.05 were identified from the comparison mad with the four stages of gastric cancer, based on two-way anova test (subtypes and stages) after false discovery rate correction (Bonferroni method). The expression levels (copy/ml) were analyzed based on the log2 scale. For each miRNA, the significant levels for the alternations between stage 1 and 2, stage 2 and 3, stage 3 and 4 were calculated based on anova test and Bonferroni adjustment to compensate for multiple comparisons. *: p-value < 0.05; **: p-value < 0.01; ***: p-value < 0.001. A miRNA was considered up-regulated if its expression level was higher in the later stage.

	anova p<0.01	Group	change between stage 2 and stage 1	change between stage 3 and stage 2	change between stage 4 and stage 3	change between normal and all cancer
hsa-miR-27a-3p	0.0008	I	-	down**	up***	No change
hsa-miR-1280	0.0007	C	up***	-	-	up-regulated
hsa-miR-579	0.0067	A	-	-	-	up-regulated
hsa-miR-150-5p	0.0035	F	-	-	-	No change
hsa-miR-29c-5p	0.0010	B	up***	-	-	up-regulated
hsa-miR-186-5p	0.0087	H	up**	down**	-	up-regulated
hsa-miR-338-5p	0.0072	L	-	-	up**	up-regulated
hsa-miR-362-5p	0.0014	B	up**	-	-	No change
hsa-miR-197-3p	0.0000	B	up***	-	-	up-regulated
hsa-miR-221-3p	0.0000	B	up***	-	-	No change
hsa-miR-501-3p	0.0000	C	up***	-	down*	No change
hsa-miR-181a-2-3p	0.0072	G	up**	-	-	No change
hsa-miR-598	0.0000	A	up**	-	-	up-regulated
hsa-miR-320b	0.0014	C	up**	-	-	No change
hsa-miR-328	0.0000	C	up***	-	down*	No change
hsa-miR-134	0.0072	D	-	-	down***	No change
hsa-miR-21-5p	0.0000	E	-	up**	down**	up-regulated
hsa-miR-424-5p	0.0000	B	up***	-	-	up-regulated
hsa-miR-99a-5p	0.0023	G	up*	-	down**	down-regulated
hsa-miR-18a-3p	0.0016	B	up***	-	-	No change
hsa-miR-195-5p	0.0000	G	up***	-	down**	No change
hsa-miR-500a-3p	0.0000	C	up***	-	down**	up-regulated
hsa-miR-18b-5p	0.0072	C	up*	-	-	up-regulated
hsa-miR-339-3p	0.0005	C	up***	-	-	No change
hsa-miR-128	0.0000	C	up***	-	down***	up-regulated
hsa-miR-22-3p	0.0016	C	-	-	down**	No change
hsa-miR-26a-5p	0.0002	G	up**	-	down***	down-regulated
hsa-miR-29b-2-5p	0.0087	B	up*	-	-	up-regulated

hsa-miR-148a-3p	0.0029	A	-	-	-	up-regulated
hsa-miR-142-5p	0.0004	H	up**	down***	up**	up-regulated
hsa-miR-23a-3p	0.0000	B	up***	-	-	No change
hsa-miR-23c	0.0002	C	up**	-	down**	down-regulated
hsa-miR-28-3p	0.0072	K	down*	-	-	No change
hsa-miR-193b-3p	0.0029	K	down**	-	-	down-regulated
hsa-miR-320a	0.0004	J	down***	up**	-	No change
hsa-miR-15b-5p	0.0000	B	up***	down*	-	No change

Table S5. Multivariate Cox regression analysis of 12-miRNA panel and clinical covariates

Variables		**Log Hazard Ratio, ln(HR)**	***p*-value**
12-miR	Cancer vs Non-cancer	13.9	< 0.001
Age (years)	> 50 vs ≤ 50	0.66	0.04
Gender	Male vs Female	0.54	0.47
Ethnicity	Chinese vs Non-Chinese	0.17	0.64
H. pylori	Yes vs No	-0.04	0.89

So JBY, *et al. Gut* 2021; 70:829–837. doi: 10.1136/gutjnl-2020-322065

Table S6. Cross-reactivity test against other common cancers

#	Type of Cancer	Number of specimens tested	Number of specimen with high risk score based on GASTROClear
1	Esophageal	12	1
2	Liver	6	1
3	Colorectal	12	3
4	Lung	12	1
5	Breast	12	0
6	Prostate	12	0
7	Kidney	12	5
8	Bladder	12	0
	Total	**90**	**11**

So JBY, *et al. Gut* 2021; 70:829–837. doi: 10.1136/gutjnl-2020-322065

Table S7. Base-case Values and Corresponding Sensitivity Range for Variables in Cost-Effectiveness Modelling

Singaporean healthcare setup			
Variable name	**Base-case value**	**Sensitivity Range**	**Source**
Costs (USD)			
MiRNA test	30	10 – 500	Assumed
Upper -endoscopy (EGD)	493	100 - 500	National University Hospital, Singapore (NUH)
Biopsy	122	-	
Stage 1 treatment	10423	-	
Stage 2 treatment	10423	-	
Stage 3 treatment	29451	-	
Stage 4 treatment	3069	-	
Follow-up examinations	719	-	
Staging Investigation (EUS + CT+ CXR+ follow-up)	1513	-	
Probabilities			
Incidence of Gastric Cancer in Chinese Males by Age group			Report No.8, 2015. Singapore Cancer Registry [26]
50 - 54 years	0.018%		
55 - 59 years	0.029%		
60 - 64 years	0.053%		
65 - 69 years	0.098%		
70 - 74 years	0.157%		
75 years	0.187%		
Stage specific diagnosis currently Stage 1: 2 : 3 : 4	18% : 11.5% : 27.5% : 43%		
Recurrence of Gastric Cancer in successfully treated patients by stage			
Recurrence in Stage 1 patients	11%	5% - 30%	Roukos *et al.* [6]
Recurrence in Stage 2 patients	53%	30% - 60%	
Recurrence in Stage 3 patients	83%	50% - 90%	

Utility Values (disutility*)			
Stage 1	0.88 (0.28)	0.78 – 0.98	Zhou HJ *et al.* [11]
Stage 2	0.86 (0.29)	0.72 – 0.99	
Stage 3	0.77 (0.31)	0.57 – 0.97	
Stage 4	0.68 (0.08)	0.52 – 0.84	
Test Characteristics			
Endoscopy Sensitivity	93%	-	Voutilainen *et al.* [12] Hamashima et *al. [27]*
Endoscopy Specificity	100%	-	Voutilainen *et al.* [12]
miRNA Sensitivity by Stages (Stage 1:2:3:4)	63% : 75% : 89% : 93%	30% - 100%	Current Study
miRNA specificity	89%	60% - 100%	Current Study

*Disutility refers to temporary reduction in QoL during first 6 months of treatment. Note: Assumed treatments are based on observed practice in Singapore. Gastric cancer patient on diagnosis undergoes staging investigation (CT, CXR, EUS & specialist consultation). Curative treatment includes surgery (total/ partial gastrectomy) & hospital stay (12days). Stage 3 patients undergo additional chemo-radiotherapy. Follow ups include: visits (2.2/year), repeat CT, CXR (1.4/year). Palliative care includes bypass surgery (30%), endoscopic stenting (6%), palliative chemotherapy (16%) & conservative treatment (2x specialist visits) (48%) with an appropriate hospital stay (12 days - on surgery, 2.5 days on average - if no surgery is performed). *Abbreviations used: CT: Computerized Tomography; CXR: Chest X-Ray; EUS: Endoscopic Ultrasound*

REFERENCE

1 Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A*, et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002;**3**:RESEARCH0034.
2 Andersen CL, Jensen JL, Orntoft TF. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res 2004;**64**:5245-50.
3 Singapore Cancer Registry Advisory Committe and Health Promotion Board S. Trends in Cancer Incidence in Singapore 2009-2013. 2014.
4 Singapore Cancer Registry. Cancer incidence and mortality 2003 - 2012 and selected trends 1973 - 2012 in Singapore. National Registry of Disease Office, 2015.
5 Department of Statistics MoTaI, Republic of Singapore. Population Trends 2016. 2016.
6 Roukos DH, Lorenz M, Karakostas K, Paraschou P, Batsis C, Kappas AM. Pathological serosa and node-based classification accurately predicts gastric cancer recurrence risk and outcome, and determines potential and limitation of a Japanese-style extensive surgery for Western patients: a prospective with quality control 10-year follow-up study. Br J Cancer 2001;**84**:1602-9.
7 Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol 2005;**5**:13.
8 Dan YY, So JB, Yeoh KG. Endoscopic screening for gastric cancer. Clin Gastroenterol Hepatol 2006;**4**:709-16.
9 Zhou HJ, Dan YY, Naidoo N, Li SC, Yeoh KG. A cost-effectiveness analysis evaluating endoscopic surveillance for gastric cancer for populations with low to intermediate risk. PLoS ONE 2013;**8**:e83959.
10 Lee S, Jun JK, Suh M, Park B, Noh DK, Jung K-W*, et al.* Gastric Cancer Screening Uptake Trends in Korea: Results for the National Cancer Screening Program From 2002 to 2011: A Prospective Cross-Sectional Study. Medicine 2015;**94**:e533.
11 Zhou HJ, So JB, Yong WP, Luo N, Zhu F, Naidoo N*, et al.* Validation of the functional assessment of cancer therapy-gastric module for the Chinese population. Health and quality of life outcomes 2012;**10**:145.
12 Voutilainen ME, Juhola MT. Evaluation of the diagnostic accuracy of gastroscopy to detect gastric tumours: clinicopathological features and prognosis of patients with gastric cancer missed on endoscopy. European journal of gastroenterology & hepatology 2005;**17**:1345-9.
13 Li C, Li J, Cai Q, Qiu Q, Yan M, Liu B*, et al.* MiRNA-199a-3p: A potential circulating diagnostic biomarker for early gastric cancer. Journal of surgical oncology 2013;**108**:89-92.
14 Gorur A, Balci Fidanci S, Dogruer Unal N, Ayaz L, Akbayir S, Yildirim Yaroglu H*, et al.* Determination of plasma microRNA for early detection of gastric cancer. Molecular biology reports 2013;**40**:2091-6.
15 Cai H, Yuan Y, Hao Y-F, Guo T-K, Wei X, Zhang Y-M. Plasma microRNAs serve as novel potential biomarkers for early detection of gastric cancer. Medical oncology (Northwood, London, England) 2013;**30**:452.
16 Cui M-H, Hou X-L, Lei X-Y, Mu F-H, Yang G-B, Yue L*, et al.* Upregulation of microRNA 181c expression in gastric cancer tissues and plasma. Asian Pacific journal of cancer prevention : APJCP 2013;**14**:3063-6.
17 Li C, Li J, Cai Q, Qiu Q, Yan M, Liu B*, et al.* miRNA-199a-3p in plasma as a potential diagnostic biomarker for gastric cancer. Annals of surgical oncology 2013;**20 Suppl 3**:405.
18 Song M-y, Pan K-f, Su H-j, Zhang L, Ma J-l, Li J-y*, et al.* Identification of serum microRNAs as novel non-invasive biomarkers for early detection of gastric cancer. PloS one 2012;**7**.

19 Li B-s, Zhao Y-l, Guo G, Li W, Zhu E-d, Luo X*, et al.* Plasma microRNAs, miR-223, miR-21 and miR-218, as novel potential biomarkers for gastric cancer detection. PloS one 2012;**7**.
20 Valladares-Ayerbes M, Reboredo M, Medina-Villaamil V, Iglesias-Díaz P, Lorenzo-Patiño M, Haz M*, et al.* Circulating miR-200c as a diagnostic and prognostic biomarker for gastric cancer. Journal of translational medicine 2012;**10**:186.
21 Zhang W-H, Gui J-H, Wang C-Z, Chang Q, Xu S-P, Cai C-H*, et al.* The identification of miR-375 as a potential biomarker in distal gastric adenocarcinoma. Oncology research 2012;**20**:139-47.
22 Lo SS, Hung PS, Chen JH, Tu HF, Fang WL, Chen CY*, et al.* Overexpression of miR-370 and downregulation of its novel target TGFβ-RII contribute to the progression of gastric carcinoma. Oncogene 2012;**31**:226-37.
23 Tsujiura M, Ichikawa D, Komatsu S, Shiozaki A, Takeshita H, Kosuga T*, et al.* Circulating microRNAs in plasma of patients with gastric cancers. Br J Cancer 2010;**102**:1174-9.
24 Liu R, Zhang C, Hu Z, Li G, Wang C, Yang C*, et al.* A five-microRNA signature identified from genome-wide serum microRNA expression profiling serves as a fingerprint for gastric cancer diagnosis. European journal of cancer (Oxford, England : 1990) 2011;**47**:784-91.
25 Liu H, Zhu L, Liu B, Yang L, Meng X, Zhang W*, et al.* Genome-wide microRNA profiles identify miR-378 as a serum biomarker for early detection of gastric cancer. Cancer Lett 2012;**316**:196-203.
26 Office NRoD. Cancer Incidence and Mortality 2003-12 and Selected Trends 1973-2012 In Singapore. Singapore Cancer Registry, Report No. 8. 2015.
27 Hamashima C, Okamoto M, Shabana M, Osaki Y, Kishimoto T. Sensitivity of endoscopic screening for gastric cancer by the incidence method. International journal of cancer Journal international du cancer 2013;**133**:653-9.