

Keywords: biomarkers; metabolomics; screening; gastric cancer; nuclear magnetic resonance

¹H-NMR urinary metabolomic profiling for diagnosis of gastric cancer

Angela W Chan¹, Pascal Mercier², Daniel Schiller³, Robert Bailey⁴, Sarah Robbins⁴, Dean T Eurich⁵, Michael B Sawyer^{*6} and David Broadhurst⁷

¹Department of Surgery, University of Alberta Hospital, 8440-112 Street, Edmonton, AB T6G 2B7, Canada; ²Department of Biochemistry, NANUC, University of Alberta, Edmonton, AB T6G 2M8, Canada; ³Department of Surgery, Royal Alexandra Hospital, 10240 Kingsway Avenue NW, Edmonton, AB T5H 3V9, Canada; ⁴Department of Medicine, Division of Gastroenterology, HYS Medical Centre, 310-11010 101 Street NW, Edmonton, AB T5H 4B9, Canada; ⁵School of Public Health, 2-040 Li Ka Shing Centre for Health Research Innovation, University of Alberta, Edmonton, AB T6G 2E6, Canada; ⁶Department of Oncology, Cross Cancer Institute, 11560 University Avenue, Edmonton, AB T6G 1Z2, Canada and ⁷Department of Medicine, 4126A Katz Group Centre for Pharmacy & Health, University of Alberta, Edmonton, AB T6G 2E1, Canada

Background: Metabolomics has shown promise in gastric cancer (GC) detection. This research sought to identify whether GC has a unique urinary metabolomic profile compared with benign gastric disease (BN) and healthy (HE) patients.

Methods: Urine from 43 GC, 40 BN, and 40 matched HE patients was analysed using ¹H nuclear magnetic resonance (¹H-NMR) spectroscopy, generating 77 reproducible metabolites (QC-RSD <25%). Univariate and multivariate (MVA) statistics were employed. A parsimonious biomarker profile of GC vs HE was investigated using LASSO regularised logistic regression (LASSO-LR). Model performance was assessed using Receiver Operating Characteristic (ROC) curves.

Results: GC displayed a clear discriminatory biomarker profile; the BN profile overlapped with GC and HE. LASSO-LR identified three discriminatory metabolites: 2-hydroxyisobutyrate, 3-indoxylsulfate, and alanine, which produced a discriminatory model with an area under the ROC of 0.95.

Conclusions: GC patients have a distinct urinary metabolite profile. This study shows clinical potential for metabolic profiling for early GC diagnosis.

Gastric adenocarcinoma (GC) is the fifth most common cancer worldwide and the third most deadly. Approximately one million people are diagnosed worldwide yearly, and there is a 70% mortality rate (Worldwide Cancer Research Fund, 2012; Cancer Research UK, 2014). GC is often diagnosed late, as non-specific symptoms, such as dyspepsia, resemble benign (BN) causes such as gastritis. In spite of this, cancers identified early have a moderate chance of cure. The 5-year survival rate of Stage IA tumours is 71% and Stage IB tumours is 57% (American Cancer Society, 2015). This highlights the importance of appropriate screening in higher-risk populations.

Metabolomics is the study of low-molecular weight chemicals (<1500 Da) in a biological system. It is the most downstream of

the 'omics' sciences (Genomics, Transcriptomics, Proteomics, etc.), and is thus considered closest to an organism's phenotype (Dunn *et al*, 2011). Previous studies show that GC cells preferentially convert glucose into lactate even in the presence of sufficient oxygen (Warburg effect) (Hirayama *et al*, 2009; Cai *et al*, 2010; Hu *et al*, 2011; Aa *et al*, 2012). Citrate is one metabolite with connections to apoptotic pathways in GC (Lu *et al*, 2011). Certain nucleic acids are overexpressed in GC (hypoxanthine, uridine, guanosine), indicating active replication (Hirayama *et al*, 2009; Hu *et al*, 2011; Yu *et al*, 2011).

Identification of a distinct urinary metabolomic profile for GC could offer a non-invasive, cost effective, efficient, and reasonably accurate modality towards accurate diagnoses. The study described

*Correspondence: Dr MB Sawyer; E-mail: michael.sawyer@albertahealthservices.ca

Received 4 May 2015; revised 10 October 2015; accepted 28 October 2015; published online 8 December 2015

© 2016 Cancer Research UK. All rights reserved 0007–0920/16

herein provides a preliminary investigation of the ability for ¹hydrogen nuclear magnetic resonance (¹H-NMR) spectroscopy to discriminate between urine samples collected from GC, healthy (HE), and benign gastric disease (BN) patients.

MATERIALS AND METHODS

Patient selection. Midstream urine samples were collected from 43 GC, 40 BN, and 40 HE patients from January 2009 to December 2014 from three hospitals in Edmonton, Canada. GC samples were collected prior to chemoradiotherapy and surgery. All patients provided written informed consent. Ethics approval was obtained from the Health Research Ethics Board at the University of Alberta.

Inclusion criteria for cancer patients were: biopsy-confirmed diagnosis of GC, age ≥ 18 years, and no metastases on their staging computed tomography scans. BN patients had to experience gastrointestinal symptoms (such as haematemesis or epigastric discomfort) and must have endoscopic evidence within the past 6 months of consent that symptoms were not due to a malignant cause. BN patients had the following conditions: gastritis, gastro-oesophageal reflux disease (GORD), portal hypertensive gastropathy, varices, gastritis, ulcers, and polyps. HE controls had no declared history of cancer and no gastrointestinal symptoms. Groups were matched on age, gender, and BMI.

Exclusion criteria included: breastfeeding, pregnancy, significant cardiac disease with New York Heart Association \geq Class II, systemic infection, prior cancer, and glomerular filtration rate $< 30 \text{ ml min}^{-1}$.

Sample collection and NMR spectroscopy. Within 2 h of collection, one ml aliquots of urine mixed with 50 μl of 0.42% sodium azide preservative were prepared and biobanked at -80°C . All one-dimensional (1D) ¹H-NMR spectra were acquired at Canada's National High Field Nuclear Magnetic Resonance Centre using a 600-MHz Varian Inova spectrometer (Agilent Inc., Palo Alto, CA, USA). Sample preparation and NMR analysis followed the standard protocols outlined in Supplementary File.

Data modelling and statistical analysis. Following standard data cleaning protocols, 77 metabolite concentrations were reproducibly detected by NMR platform. For each metabolite, pairwise comparisons of GC vs HE and BN vs HE were tested using the non-parametric Mann–Whitney *U*-test. Correction for multiple comparisons was performed using Benjamini and Hochberg method (Benjamini, 1995).

Exploratory multivariate statistical analysis in the form of Partial Least Squares discriminant analysis (PLS-DA) and Orthogonal Partial Least Squares discriminant analysis (O-PLS-DA) were used to uncover any latent correlated structure in the data (Eriksson *et al*, 2013). Logistic regression optimised by LASSO regularisation (LASSO-LR) was then performed to derive a parsimonious discriminant GC vs HE biomarker model. Statistical analyses were performed using SIMCA (version 13, Umetrics, Umea, Sweden), Matlab scripting language (MathWorks Inc., Natick, MA, USA), and STATA Version 13 (StataCorp LP, College Station, TX, USA).

RESULTS

Patient characteristics. Baseline patient and tumour characteristics are listed in Table 1. To compare univariate statistical results from two arms of this study (GC vs HE and BN vs HE), a bi-plot of log median fold change for metabolites significant in either comparison was constructed (Figure 1). *P*-values, *q*-values, median

Table 1. Baseline characteristics of the study subjects and tumour

Characteristic	BN	GC	HE
Number of patients	40	43	40
Mean age (s.d.), years	63.1 (9.0)	65.2 (12.0)	63.2 (8.8)
Gender (male/female)	19/21	28/15	23/17
Mean BMI (s.d.), kg m^{-2}	29.5 (6.4)	27.6 (6.9)	27.7 (4.7)
<i>Helicobacter pylori</i> status (on biopsy or urea breath test)			
Positive/negative/unknown	3/21/16	7/26/10	—
Benign condition			
Gastritis only	13 (32.5%)	—	—
Ulcer only	4 (10.0%)	—	—
Gastritis and ulcer	1 (2.5%)	—	—
Gastritis and portal hypertensive gastropathy (PHG)	1 (2.5%)	—	—
PHG	9 (22.5%)	—	—
Gastro-oesophageal reflux disease (GORD)	3 (7.5%)	—	—
Varices	1 (2.5%)	—	—
Polyps	5 (12.5%)	—	—
Reactive gastropathy	1 (2.5%)	—	—
Normal scope with GI symptoms	2 (5.0%)	—	—
Overall TNM stage			
Ia/b	—	3/3	—
Ila/b	—	8/3	—
IIla/b/c	—	2/5/3	—
IV	—	14	—
Unknown	—	2	—
Tumour location			
GE junction/cardia/fundus/body/antrum/pylorus	—	6/1/4/15/16/1	—
Lauren histological class			
Diffuse/intestinal/mixed/not specified	—	15/16/3/9	—
Grade (differentiation)			
Well/moderate/moderate to poor/poor/not reported	—	3/8/5/29/3	—
Resectable/not resectable			
Neoadjuvant (yes/no)	—	10/18	—
Adjuvant (yes/no)	—	18/10	—

Abbreviations: BMI = body mass index; BN, benign gastric disease; GC = gastric cancer; GE = gastro-oesophageal; GI = gastrointestinal; HE = healthy; TNM = tumour, node, metastasis.

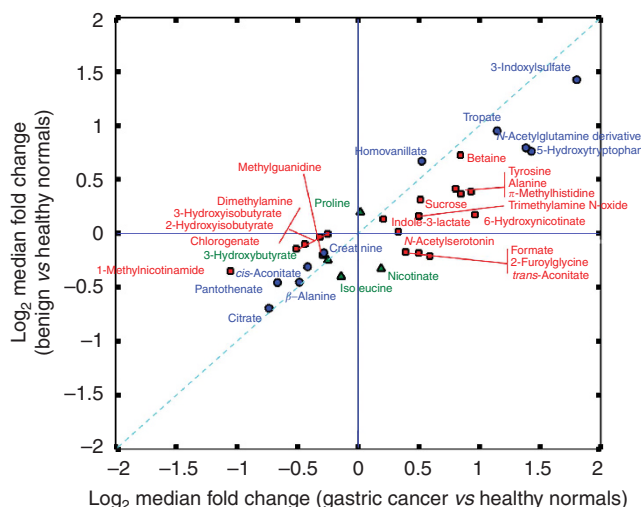


Figure 1. Biplot of log₂ median fold change for metabolites in GC vs HE and BN vs HE models. Blue circles represent metabolites significantly changed in both models; red squares, significantly changed in GC vs HE only; green triangles, significantly changed in BN vs HE only.

concentrations, and median-fold differences for each pairwise comparison are reported in Supplementary Table S1. A detailed discussion of the PLS-DA and O-PLS-DA models are provided in Supplementary Files. These results reflect those of the univariate statistics. Of particular interest were nine metabolites, which had high VIP scores in the GC vs HE OPLS model but low VIP scores in the BN vs HE OPLS model (Supplementary Table S2): sucrose, dimethylamine, 1-methylnicotinamide, 2-furoylglycine, N-acetylserotonin, trans-aconitate, alanine, formate, and serotonin.

LASSO-LR produced an optimal GC vs HE model using just three metabolites: 2-hydroxyisobutyrate (2-HIB), 3-indoxylsulfate (3-IS), and alanine (A). This resulted in the following diagnostic regression model:

$$(P) = 15.0 - 7.7 \times \log(2\text{-HIB}) + 5.2 \times \log(3\text{-IS}) - 6.1 \times \log(A)$$

The corresponding ROC curve had an AUC of 0.95 (95% CI: 0.86–0.99) (Figure 2A). For a fixed specificity of 80%, the corresponding sensitivity for predicting GC was 95% (95% CI: 0.86–0.99). According to this specificity, if the predicted score, P , for a given individual is >0.3 the diagnosis would be 'GC'; otherwise if $P < 0.3$, 'not GC'. Figure 2B shows a frequency histogram for three disease classifications grouped by the LASSO-

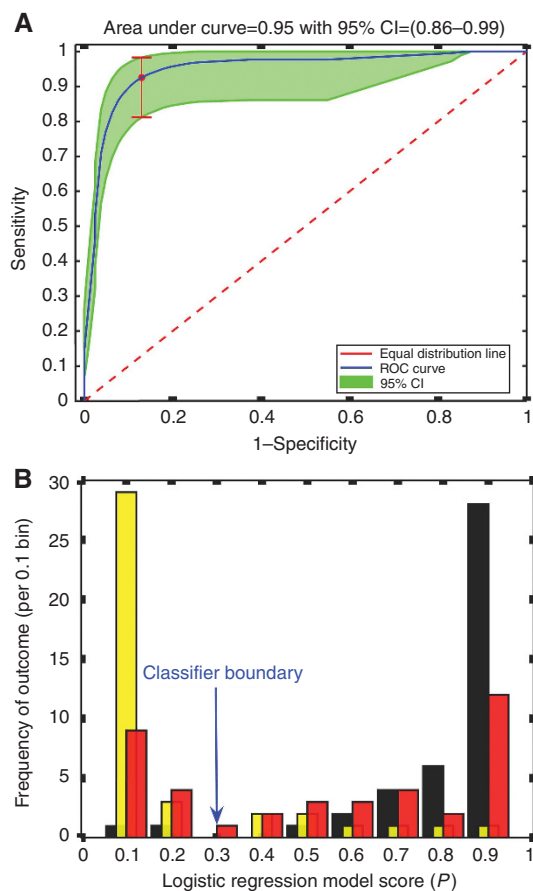


Figure 2. Three-metabolite logistic regression model. **(A)** Receiver Operating Characteristic (ROC) curve for GC vs HE comparison based on three-metabolite model. Area under the curve (AUC) is 0.95 (95% CI = 0.86–0.99). For a fixed specificity of 80%, the sensitivity is 95% (95% CI = 0.85–1.00). **(B)** Frequency histogram for logistic regression model scores. Yellow bars represent HE patients; red, BN patients; and black, GC patients. The number (frequency) of patients with each score is depicted by the height of the bars. Scores closer to 1 indicate a high probability of GC; close to 0 indicates high probability of HE. Cutoff boundary is score 0.3. Above 0.3, classified as GC; below, not GC.

LR model score. BN samples are split into two broad distributions: half of BN patients classified with GC, and the other half with HE.

DISCUSSION

GC is a highly morbid and fatal disease. Diagnosis of GC is often delayed. The present study used 1D $^1\text{H-NMR}$ spectroscopy to characterise a urinary metabolic profile of GC that is distinct from HE and a subpopulation of BN patients.

Five to seven percent of skeletal muscle is composed of alanine, an endogenous amino acid (Felig *et al*, 1978). During fasting, muscle protein is catabolised to release alanine for liver gluconeogenesis. Similar to previous studies (Hirayama *et al*, 2009; Chen *et al*, 2010), alanine concentration increased from HE to GC. Elevated alanine levels in GC patients' urine compared with HE show that alanine may be a biomarker of muscle wasting but not necessarily a specific biomarker of the disease itself.

In rats with chemically induced gastric lesions (ulcers, erosions), treatment with 1-methylnicotinamide inhibited gastric acid secretion and increased mucosal blood flow and healing (Brzozowski *et al*, 2008). Diminished levels of 1-methylnicotinamide in both BN and GC groups suggest loss of this mucosal protective mechanism. Where mucosa is ulcerated or eroded, sucrose can penetrate more easily into the bloodstream and be excreted into the urine (Sutherland *et al*, 1994). Our study shows significant sucrose elevations in both BN and GC groups compared with HE. Perhaps this is due to the increased permeability of damaged mucosa in GC and BN patients.

Creatinine, a waste product of muscle metabolism, is excreted by the kidneys (Eisner *et al*, 2011). The amount of creatinine in urine is related to muscle mass (Swaminathan *et al*, 2000). Cachectic patients have lower total body skeletal muscle mass and therefore lower levels of urinary creatinine. This phenomenon was consistent with our results.

Citrate is an intermediate of the Krebs' cycle. An *in vitro* experiment showed that citrate induced apoptosis in two GC cell lines in a dose-dependent manner (Lu *et al*, 2011). In our study, citrate was downregulated in GC patients, suggesting an ability of GC to escape regular programmed cell death.

The distinction between BN and either GC or HE was less clear using the multiclass PLS model (Supplementary Figure S2). BN conditions that clustered more frequently with GC include: ulcers, GORD, and gastritis. These observations fit with Correa's hypothesis (Correa, 1988). He delineated a preneoplastic cascade from healthy to chronic atrophic gastritis and eventually to cancer. Patients with chronic gastritis are farther on the preneoplastic cascade than early gastritis patients, so their phenotypes and metabolomic signatures more likely resemble GC than HE.

This observational study has limitations. We enrolled a pragmatic sample size of roughly 40 patients in each group. A small sample size limits the power to detect a difference, and conversely, differences detected may be spurious. This experiment matched patients on three common confounders – age, sex, and BMI, but as it is an observational design, only known confounders can be controlled. Other confounders in this experiment include: medications, smoking, and *Helicobacter pylori* status.

This study shows clinical potential for metabolic profiling, although numerous steps are required to move this test into the clinic. Should the three-metabolite model be successfully validated, then a point-of-care diagnostic could be developed such as a simple dipstick or laboratory assay. Alternatively, if the complete metabolite profile needs to be done, then this assay could be performed at a centralised laboratory with samples being collected and processed in the periphery.

ACKNOWLEDGEMENTS

We thank the following individuals and organisations for their contributions to this manuscript: laboratory technicians Michelle Kuzma and Delores Mowles; Rose Cornand, for assistance with patient consent and urine collection at the Royal Alexandra Hospital; Gastroenterologists (Dr Clarence Wong), surgeons (Dr Klaus Buttenschoen, Dr Cliff Sample, Dr Erika Haase, Dr Gordon Lees) and their respective administrative staff for referring study patients; endoscopy suite staff at the Royal Alexandra and University of Alberta Hospitals for use of their facilities; financial support: Edmonton Civic Employees Charitable Assistance Fund, Alberta Innovates Technology Futures Graduate Student Scholarship, Dr D Schiller Academic Enrichment Fund, Queen Elizabeth II Graduate Scholarship; Dr Dean Eurich holds a Canada Research Chair from Government of Canada and Alberta Heritage Foundation for Medical Research (AHFMR) Population Health Investigator Award from Alberta Innovates Health Solutions (AIHS).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Aa J, Yu L, Sun M, Liu L, Li M, Cao B, Shi J, Xu J, Cheng L, Zhou J, Zheng T, Wang X, Zhao C, Gu R, Zhang F, Shi R, Wang G (2012) Metabolic features of the tumor microenvironment of gastric cancer and the link to the systemic macroenvironment. *Metabolomics* **8**(1): 164–173.
- American Cancer Society (2015) *Survival rates for stomach cancer, by stage* [Online]. American Cancer Society. Available at <http://www.cancer.org/cancer/stomachcancer/detailedguide/stomach-cancer-survival-rates> (accessed 2015).
- Benjamini YHY (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol* **57**(1): 289–300.
- Brzozowski T, Konturek PC, Chlopicki S, Sliwowski Z, Pawlik M, Ptak-Belowska A, Kwiecien S, Drozdowicz D, Pajdo R, Slonimska E, Konturek SJ, Pawlik WW (2008) Therapeutic potential of 1-methylnicotinamide against acute gastric lesions induced by stress: role of endogenous prostacyclin and sensory nerves. *J Pharmacol Exp Ther* **326**(1): 105–116.
- Cai Z, Zhao JS, Li JJ, Peng DN, Wang XY, Chen TL, Qiu YP, Chen PP, Li WJ, Xu LY, Li EM, Tam JP, Qi RZ, Jia W, Xie D (2010) A combined proteomics and metabolomics profiling of gastric cardia cancer reveals characteristic dysregulations in glucose metabolism. *Mol Cell Proteomics* **9**(12): 2617–2628.
- Cancer Research UK (2014) *Stomach (gastric) cancer* [Online]. Cancer Research UK. Available at <http://www.cancerresearchuk.org/cancer-info/cancerstats/incidence/> (accessed web page 2014).
- Chen JL, Tang HQ, Hu JD, Fan J, Hong J, Gu JZ (2010) Metabolomics of gastric cancer metastasis detected by gas chromatography and mass spectrometry. *World J Gastroenterol* **16**(46): 5874–5880.
- Correa P (1988) A human model of gastric carcinogenesis. *Cancer Res* **48**(13): 3554–3560.
- Dunn WB, Broadhurst DI, Atherton HJ, Goodacre R, Griffin JL (2011) Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chem Soc Rev* **40**(1): 387–426.
- Eisner R, Stretch C, Eastman T, Xia J, Hau D, Damaraju S, Greiner R, Wishart D, Baracos V (2011) Learning to predict cancer-associated skeletal muscle wasting from 1h-NMR profiles of urinary metabolites. *Metabolomics* **7**(1): 25–34.
- Eriksson L, Byrne T, Johansson E, Trygg J, Vikstrom C (2013) *Multi- and Megavariate Data Analysis, Basic Principles and Applications*. MKS Umetrics AB: Malmö, Sweden.
- Felig P, Pozefsky T, Marliss E, Cahill Jr GF (1978) Alanine: key role in gluconeogenesis. *Science* **167**(3920): 1003–1004.
- Hirayama A, Kami K, Sugimoto M, Sugawara M, Toki N, Onozuka H, Kinoshita T, Saito N, Ochiai A, Tomita M, Esumi H, Soga T (2009) Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry. *Cancer Res* **69**(11): 4918–4925.
- Hu JD, Tang HQ, Zhang Q, Fan J, Hong J, Gu JZ, Chen JL (2011) Prediction of gastric cancer metastasis through urinary metabolomic investigation using gc/ms. *World J Gastroenterol* **17**(6): 727–734.
- Lu Y, Zhang X, Zhang H, Lan J, Huang G, Varin E, Lincet H, Poulain L, Icard P (2011) Citrate induces apoptotic cell death: a promising way to treat gastric carcinoma? *Anticancer Res* **31**(3): 797–805.
- Sutherland LR, Verhoef M, Wallace JL, Van Rosendaal G, Crutcher R, Meddings JB (1994) A simple, non-invasive marker of gastric damage: sucrose permeability. *Lancet* **343**(8904): 998–1000.
- Swaminathan R, Major P, Snieder H, Spector T (2000) Serum creatinine and fat-free mass (lean body mass). *Clin Chem* **46**(10): 1695–1696.
- Worldwide Cancer Research Fund (2012) *Stomach cancer statistics* [Online]. Worldwide Cancer Research Fund. Available at <http://www.wcrf.org/int/cancer-facts-figures/data-specific-cancers/stomach-cancer-statistics> (accessed web page 2014).
- Yu L, Aa J, Xu J, Sun M, Qian S, Cheng L, Yang S, Shi R (2011) Metabolomic phenotype of gastric cancer and precancerous stages based on gas chromatography time-of-flight mass spectrometry. *J Gastroenterol Hepatol* **26**(8): 1290–1297.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 4.0 Unported License.

Supplementary Information accompanies this paper on British Journal of Cancer website (<http://www.nature.com/bjc>)