Expression Profiles of Loneliness-associated Genes for Survival Prediction in Cancer Patients

Liang-Fu You1*, Jia-Rong Yeh2, Mu-Chun Su1

Abstract

Influence of loneliness on human survival has been established epidemiologically, but genomic research remains undeveloped. We identified 34 loneliness-associated genes which were statistically significant for high-lonely and low-lonely individuals. With the univariate Cox proportional hazards regression model, we obtained corresponding regression coefficients for loneliness-associated genes for individual cancer patients. Furthermore, risk scores could be generated with the combination of gene expression level multiplied by corresponding regression coefficients of loneliness-associated genes. We verified that high-risk score cancer patients had shorter mean survival time than their low-risk score counterparts. Then we validated the loneliness-associated gene signature in three independent brain cancer cohorts with Kaplan-Meier survival curves (n=77, 85 and 191), significantly separable by log-rank test with hazard ratios (HR) >1 and p-values <0.0001 (HR=2.94, 3.82, and 1.78). Moreover, we validated the loneliness-associated gene signature in bone cancer (HR=5.10, p-value=4.69e-3), lung cancer (HR=2.86, p-value=4.71e-5), ovarian cancer (HR=1.97, p-value=3.11e-5), and leukemia (HR=2.06, p-value=1.79e-4) cohorts. The last lymphoma cohort proved to have an HR=3.50, p-value=1.15e-7. Loneliness-associated genes had good survival prediction for cancer patients, especially bone cancer patients. Our study provided the first indication that expression of loneliness-associated genes are related to survival time of cancer patients.

Keywords: Loneliness-associated genes - expression profile - hazards ratios - cancer patients - survival

Introduction

The influence of loneliness (social isolation) on human health has been verified by epidemiologically, histologically, and genomically. Liillberg et al., indicated that a twofold increase in breast cancer risk owing to disruption of marriage, e.g. divorce, separation or death of a spouse, by epidemiological data of 10808 women (Lillberg et al., 2003). Holwerda et al., found that mortality hazard ratio for feeling of loneliness was 1.3 in men, 1.04 in women with a 10-year follow-up of 4,004 older persons (Holwerda et al., 2012). In 2007, Cole et al., first provided that subjective social isolation was associated with human genome-wide transcriptional activity, including inhibiting anti-inflammatory transcription and promoting pro-inflammatory pathways (Cole et al., 2007). Furthermore, Lutgendorf et al. (2009) validated that high depression and low social support was relative to increased activity of β-adrenergic transcription control, which promotes tumor progression in ovarian cancer. With gene expression of a genome, survival prediction in cancer patients was improved over histologic grades. In 2004, Freije et al., proved that gene expression of genome for clustering patients was a more powerful survival predictor than histologic grades for brain cancer (Freije et al., 2004). Metzeler et al., derived 86 probe-sets of gene expression to predict survival for leukemia (Metzeler et al., 2008). In breast cancer, Motakis et al., used a couple of genes for survival prediction (Motakis et al., 2009). Two articles derived the survival significant genes and implemented Cox proportional hazards model to predict survival between high-risk and low-risk patients in non-small cell lung cancer (Hsu et al., 2009; Hou et al., 2010). Yoshihara et al., derived 88 survival significant genes for predicting survival in ovarian cancer (Yoshihara et al., 2010). Our study focused on loneliness-associated genes for survival prediction in different kinds of cancer patients, and the results showed that loneliness-associated genes influenced the survival time of cancer patients. We concluded that the difference of loneliness-associated genes are statistically significant in high-lonely and low-lonely individuals, and with Cox proportional hazards regression model, we got the risk scores of each patients with the combination of gene expression level multiplied the corresponding regression coefficients of the loneliness-associated genes. Cancer patients were clustered into the

1Department of Computer Science and Information Engineering, National Central University, 2Research Center for Adaptive Data Analysis, National Central University, ChungLi, Taiwan *For correspondence: davidyou7@gmail.com
We verified that the high-risk scores of cancer patients have shorter mean survival time than the low-risk scores of cancer patients; the Kaplan-Meier survival curves showed that high-risk versus low-risk cancer patients were significantly separable and had hazard ratio (HR) > 1, and p-value < 0.0001 with a log-rank test.

Materials and Methods

Schematic diagram of survival prediction

The strategy of our approach to predict the survival time of cancer patients with loneliness-associated genes was illustrated in the schematic diagram of Figure 1. We figured out the loneliness-associated genes by differential expression between the high and the low lonely groups by the Student’s t-test. The loneliness-associated genes were used to calculate the risk score for each cancer patient for survival prediction, which is detailed in the following part of the survival analysis.

Loneliness-associated genes

Based on Affymetrix microarray of GSE7148 gene expression profiles, we wanted to figure out the loneliness-associated genes by finding the genes with differential expression between the high and the low lonely groups by the Student’s t-test with p-value < 0.005.

Participants of GSE7148 came from 230 individuals participating in the Chicago health aging and social relations study (CHASRS) and they were 50-67 years old (Cacioppo et al., 2006). CHASRS was a 5-year cohort analysis, 10 individuals experiencing high level of subjective social isolation were identified by scores in the top 15% of the loneliness scale during first 3 years, and 10 individuals who consistently scored in the bottom 15% during first 3 years. Then gene expression profiles were carried out on peripheral blood leukocytes from these 10 individuals in the top 15% of the loneliness scale and these 10 individuals in the bottom 15%. Then GSE7148 got 14 samples of 7 high lonely participants used for high lonely group and 7 low lonely participants used for low lonely group after excluding the hesitative samples. High lonely group reported higher level of loneliness, depressive symptoms, perceived stress, and hostility. And high lonely group differed from low lonely group not only on a lower household income of mean 425,000 versus 916,000 yearly but also less married percentage in marital status. The characteristics of psychological participants of high lonely and low lonely groups were listed in Table 1 (Cole et al., 2007). Among the table, the revised UCLA loneliness scale was used for measuring the loneliness that had become the most common instrument used by researchers in assessing feelings of loneliness since 1980 (Russell et al., 1980) And the center for epidemiological studies depression scale (CESD) was a 20-item measure that assesses depressive symptoms, and scores of 16 or higher have been associated with clinical depression. (Gonzalez-Fortezza et al., 2011) Alternatively, the perceived stress scale was a 14-item measure designed to assess the degree to which situations in one’s life over the past month were perceived as stressful (Cohen et al., 1983). Moreover, high scores on the Cook-Medley hostility scale have been associated with enhanced risk for physical disorders (Contrada and Jussim 1992).

We used the gene expression profiles of GSE7148 to find loneliness-associated genes between high lonely and low lonely people that were carried out on peripheral blood leukocytes from 14 healthy adults, 50-67 years old.

Survival analysis

The association between survival and each gene expression profile from microarray was evaluated by univariate Cox proportional hazard regression analysis. More specifically, a patient’s risk score was calculated as the sum of the levels of expression of each gene measured.

Table 1. Characteristics of Psychological Participants

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>High-lonely</th>
<th>Low-lonely</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCLA Loneliness (mean±SD*)</td>
<td>46.0±5.6</td>
<td>29.9±5.1</td>
</tr>
<tr>
<td>Depressive symptoms (CESD mean±SD)</td>
<td>15.3±11.9</td>
<td>1.9±2.8</td>
</tr>
<tr>
<td>Stress (Perceived Stress Scale mean±SD)</td>
<td>15.8±5.3</td>
<td>7.5±6.6</td>
</tr>
<tr>
<td>Hostility (Cook-Medley Hostility Inventory mean±SD)</td>
<td>17.2±8.2</td>
<td>11.1±7.0</td>
</tr>
</tbody>
</table>

Table 1. Characteristics of Psychological Participants


Figure 1. Schematic Diagram of the Survival Prediction for Cancer Patients with Loneliness-associated Genes
Expression Profiles of Loneliness-associated Genes for Survival Prediction in Cancer Patients

Table 2. Cox Proportional Hazards Regression Survival Analysis for Cancer Patients

<table>
<thead>
<tr>
<th>Cancers</th>
<th>Patients</th>
<th>Datasets</th>
<th>High-risk patients</th>
<th>Low-risk patients</th>
<th>Hazard Ratio*</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain cancer (glioma)</td>
<td>77</td>
<td>GSE4271</td>
<td>19.7</td>
<td>45.1</td>
<td>2.94</td>
<td>4.58e-6</td>
</tr>
<tr>
<td>Brain cancer (glioma)</td>
<td>85</td>
<td>GSE4412</td>
<td>9.0</td>
<td>30.2</td>
<td>3.82</td>
<td>1.37e-8</td>
</tr>
<tr>
<td>Brain cancer (glioma)</td>
<td>191</td>
<td>GSE13041</td>
<td>14.6</td>
<td>24.1</td>
<td>1.78</td>
<td>7.71e-5</td>
</tr>
<tr>
<td>Bone cancer</td>
<td>34</td>
<td>GSE16091</td>
<td>51.0</td>
<td>74.4</td>
<td>5.10</td>
<td>4.69e-3</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>111</td>
<td>GSE3141</td>
<td>25.5</td>
<td>41.3</td>
<td>2.86</td>
<td>4.71e-5</td>
</tr>
<tr>
<td>Ovarian cancer</td>
<td>153</td>
<td>GSE26712</td>
<td>32.2</td>
<td>52.7</td>
<td>1.97</td>
<td>3.11e-5</td>
</tr>
<tr>
<td>Leukemia</td>
<td>163</td>
<td>GSE12417</td>
<td>10.6</td>
<td>19.6</td>
<td>2.06</td>
<td>1.79e-4</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>159</td>
<td>GSE4475</td>
<td>29.9</td>
<td>52.8</td>
<td>3.50</td>
<td>1.15e-7</td>
</tr>
</tbody>
</table>

*The Hazard Ratio was reported for the high-risk versus the low-risk cancer patients as determined by the risk scores, and the risk score was a linear combination of the gene expression levels weighted by the corresponding Cox regression coefficients.

In our studies, we exploited the databases of Gene Expression Omnibus (GEO) GSE7148 to retrieve loneliness-associated genes with the Affymetrix microarray of platform GPL96, and we retrieved the public multiple human cancers cohorts used for validation from independent microarray profiles with different kinds of cancer patients. Three cohorts of brain cancer (glioma) datasets came from the Gene Expression Omnibus GSE4412 (Freije et al., 2004), GSE4271 (Phillips et al., 2006), and GSE13041 (Lee et al., 2008). And the bone cancer (osteosarcoma), lung cancer, ovarian cancer, leukemia, and lymphoma cohort came from Gene Expression Omnibus GSE16091 (Paoloni et al., 2009), GSE3141 (Bild et al., 2006), GSE26712 (Bonome et al., 2008), GSE12417 (Metzeler et al., 2008), and GSE4475 (Hummel et al., 2006), individually. The data sets of GEO databases were downloaded from the National Center for Biotechnology Information (NCBI) in the web site of http://www.ncbi.nlm.nih.gov/.

We then used the mathematic tool of MATLAB to conveniently complete our calculation, Microsoft software of Assess to get the corresponding genes and their gene expression profiles in different GEO datasets which were saved with Microsoft software of Excel. And the software of R was the tool to acquire the Kaplan-Meier survival curves and the log-rank test of high-risk vs. low-risk cancer patients.

Results

Loneliness-associated genes for survival prediction

Based on Affymetrix GSE7148 gene expression data, we identified the 34 loneliness-associated genes from microarray gene expression data set by finding the genes with differential expression between the high and the low lonely groups by the Student’s t-test with p-value < 0.005. The 34-gene signature of loneliness-associated genes with good survival prediction potentials was obtained. Finally, we validated the 34-gene signature of loneliness-associated was suit for survival prediction in many kinds of cancer patients by using public microarray profiles.

Survival prediction for high-risk vs. low-risk cancer patients

In Table 2 we showed that the high-risk score cancer patients have shorter mean survival time than the low-risk score cancer patients; it also shows the hazard ratio
of high-risk and low-risk cancer patients with microarray datasets.

Then we illustrated the loneliness-associated genes for survival prediction in three independent brain cancer cohorts (patients=77, 85, and 191) and bone cancer cohort (patients=34) with Kaplan-Meier survival curves within Figure 2. Kaplan-Meier survival curves of log-rank test in brain cancer cohorts were significantly separable and had hazard ratio (HR=2.94, 3.82, and 1.78), p-value <0.0001 with log-rank test. Moreover, we validated the loneliness-associated gene signature in bone cancer cohort (HR=5.10, p-value=1.79e-4), lung cancer cohort (HR=2.86, p-value=4.71e-5), ovarian cancer cohort (HR=1.97, p-value=3.11e-5), and leukemia cohort (HR=2.06, p-value=1.79e-4). The last lymphoma cohort was proved with HR=3.50, p-value=1.15e-7.

In Figure 2 we illustrated the survival distributions of three brain cancer cohorts and one bone cancer cohort; there were 77 patients in brain cancer cohort (GSE4271), 185 patients in brain cancer cohort (GSE4412), 191 patients in brain cancer cohort (GSE13041), and 34 patients in bone cancer cohort (GSE16091), separately with Kaplan-Meier survival curves.

In Figure 3 we illustrated the survival distributions of 111 patients in lung cancer cohort (GSE3141), 153 patients in ovarian cancer cohort (GSE26712), 163 patients in leukemia cohort (GSE12417), and 159 patients in lymphoma cohort (GSE4475) with Kaplan-Meier survival curves.

**Discussion**

Some studies examined the effectiveness of medicine for treating cancer patients by gene expression profiles of microarray. (Yoshino et al., 2011) Microarray datasets have also been widely used in discovering tumor types and tumor progression by comparing different stages of tumors (Privette Vinnedge et al., 2011; Yeoh et al., 2010). Moreover, some studies developed a gene expression signature for predicting overall survival of lung cancer with microarray data sets (Yoshihara et al., 2012).

Psychological states such as chronic loneliness, depression, and stress influenced cancer progression, usually because of a degradation of the human immune response (Webster Marketon and Glaser 2008; Dhabhar et al., 2012). And it was found that psychology influenced cancer progression via stimulating growth factor of cancer or tumor angiogenesis for multiple human cancers (Lutgendorf et al., 2003; Nausheen et al., 2010).

There were less scientific descriptions for alternative treatment of cancer patients, such as healing by nutrition, psychology, music, spirit, etc. (Dusek et al., 2008; Ornish et al., 2008; Dein and Pargament 2012). More especially, many cancer patients were cured by just praying in Jesus’ name for healing – these can be seen on YouTube. Healing cancer patients in that way seems incredible and differs from traditional treatments. But more and more cancer patients have been healed by alternative treatments. So we should do our best to discover such miracles with scientific evidence and let more cancer patients benefit from our scientific efforts (Hodge, 2010).

We exploited the knowledge of gene expression and survival analysis in cancer patients to show that the high-lonely people who had more negative psychology had a higher risk score. A linear combination of the gene expression values was weighted by regression coefficients; we then made use of exponential functions of this linear combination to get the risk score for the high and the low groups.

Many studies have developed useful gene signatures by examining only the gene profiling data of the patient specimens that they can access (Lu et al., 2006; Raponi et al., 2006; Beer et al., 2002).

Differing from these works, our latest approach started with an in silicon exploitation of the public domain microarray databases to derive the gene signature and concluded with validation by public independent clinical data (Hsu et al., 2009; Guo and Wan 2012).

Some biologists might doubt that loneliness-associated genes can be used for survival prediction in cancer patients. In fact our results of survival prediction showed the 34-gene signature derived from loneliness-associated genes was suit for survival prediction in cancer patients, and it was as good as the gene signature derived from other biological methods, eg. the metastatic cells approach in paper (Hsu et al., 2009).

Loneliness not only might influence cancer’s development, but also cancer progression by inducing the immune system of the human body. In such case loneliness should also influence the survival time of cancer patients, and our experiment design deriving the gene signature from loneliness factor for survival prediction was reasonable. Finally, our study validated that gene signatures derived from loneliness profiles have good survival prediction for cancer patients, which means a psychological factor of loneliness might play an important role in the survival of cancer patients.

Moreover, we used the FLink (Frequency weighted...
Links) that examined the genes most frequently linked to the biological biosystem by the web site of National Center for Biotechnology Information (NCBI) to analyze the loneliness-associated genes statistically significant different ($p<0.01$) between the high-lonely and low-lonely groups. Then we found that 16 genes of the loneliness-associated genes were involved in the human immune system, 14 genes were involved in metabolism, and 11 genes were involved in the signal transduction of human biosystem. Human immune system prevented infection and disease. Metabolic processes mainly generated energy through the consumption of nutrition in the food and released the unused materials. And signal transduction was a signaling process in which extracellular signals, such as hormones, growth factors, or temperature, elicit changes in cell state or activity. Stimulation of signal receptors lead to intracellular environment changes by activating downstream signaling process, which might impact cellular proliferation, differentiation, and survival for human beings. From the results of FLink analysis, the loneliness-associated genes were mainly associated with human immune system, metabolism, and signal transduction of human biosystem.

Negative psychology e.g. stress, depression, hostility, or loneliness may originate from trauma, pessimism, low economic income, addiction, unmarried, or social isolation (Sharma et al., 2011; Sperlich et al., 2011). Negative psychology not only yielded cancer generation, development, and progression but also reduced the survival time of cancer patients by lowering human immune functions or stimulating the signaling pathway of cancer in the human body (Conti et al., 2011; Lamkin et al., 2012).

Traditionally curing cancer patients with tumor excision, chemical medicine, and radiation treatment had many side effects for humans that wrecked the immune functions (Kirchheiner et al., 2013; Rueda-Lara, 2013). That’s the reason why many cancer patients died from just catching a cold or other infection after traditional treatments.

Alternatively, let’s focus on the importance of psychology influencing human health, and keep ourselves joyful by regular exercise, ordinarily daily routine. (Radom-Aizik et al., 2008).

More especially, let’s trust in God and read the Bible, which is the most popular and profitable book in the world, which will lead us a healthy living.

Our results showed that loneliness-associated genes influenced even the survival time of cancer patients. And we concluded that the gene signature based on loneliness approach, which is derived from microarray profiles, had good survival prediction power in cancer patients.

**Acknowledgements**

This paper was partly supported by the National Science Council, Taiwan, R.O.C, under NSC 101-2221-E-008-124-MY3, 101-2911-I-008-001, the CGH-NCU Joint Research Foundation Project number: 102 NCU-CGH A4, and the LSH-NCU Joint Research Foundation Project number: 102- NCU-LSH-102-A-018.

**References**


