Al-Salam Journal for Medical Science
Journal Homepage: http://journal.alsalam.edu.iq/index.php/ajbms
E-ISSN: 2959-5398, P-ISSN: 2958-0870

Role of speA Gene Expression in Streptococcus pyogenes of Infectious Diseases

Rawaa T. Abdulghafoor [1]*

[1]General Directorate for Education of Anbar, Ministry of Education, Anbar, Ramadi, Iraq.

*Corresponding Author: Rawaa T. Abdulghafoor

DOI: https://doi.org/10.55145/ajbms.2025.06.02.010
Received May 2025; Accepted August 2025; Available August 2025

ABSTRACT: Streptococcus pyogenes (Group A Streptococcus, GAS) is an important human pathogen that can cause a wide range of clinical conditions, ranging from mild (i.e. pharyngitis) to severe and life-threatening diseases like necrotizing fasciitis and streptococcal toxic shock syndrome. The balance of inducing mild versus invasive disease will likely be determined by several virulence factors such as the pyrogenic exotoxins. The *speA* gene, which encodes Streptococcal pyrogenic exotoxin A *SpeA* has been associated with severe invasive disease. The present study is intended to examine the expression of the *speA* gene in S. pyogenes (group A Streptococcus, GAS) strains recovered from both invasive and mild infections. Twenty clinical isolates of S. pyogenes (10 from patients with invasive infections and 10 from those having mild infections) were used. Gene expression of *speA* was examined by quantitative real-time PCR (RT-qPCR) with 16S rRNA gene as a reference. Results The level of *speA* expression were significantly greater in the invasive isolates than those from non-invasive infections and there was a strong correlation between *speA* expression and the invasiveness of individual invasive isolates. Changes in *speA* expression, as determined by unpaired statistical analysis with the use of the t-test, were highly significant ($p<0.001$). In addition, logistic regression analysis showed that *speA* gene expression was a significantly positive predictor for invasive infection. Our results suggest that *speA* expression is a potential biomarker that can aid in predicting the pathogenicity nature of S. pyogenes isolates, and may also direct future therapeutic and diagnosis strategies for eradicating these virulence determinants in invasive disease.

Keywords: Pyogenes, *SpeA* gene, Streptococcus

1. INTRODUCTION

Streptococcus pyogenes (group A Streptococcus, GAS) is one of the frailest human pathogens that causes a spectrum of diseases ranging from mild infections, such as pharyngitis, to severe life-threatening diseases, such as necrotizing fasciitis and toxic shock syndrome [1]. So far identified major virulence factors included, M protein, streptococcal pyrogenic exotoxins, and superantigens, and all of these factors contribute to the ability of pathogen to cause non-invasive (pharyngitis) and invasive disease. [2] Infections with the S. pyogenes strains can be either benign (e.g., pharyngitis) or invasive and lead to severe disease (e.g., sepsis, toxic shock) [3]. More virulent forms of infection tend to be associated with rapidly necrotising tissues and can be fatal within days [4]. Much of the disease pathogenesis is driven by virulence factors including streptolysin O which is cytolytic, and pyrogenic exotoxins which can cause systemic inflammation and shock [5]. These virulence factors enable the bacterium to evade host immunity, allowing it to persist and even transit to systemic infection [6]. Several M-types, particularly M1 and M3, have been shown to be associated with invasive infection [7]. The severe clinical manifestations, including organ failure and septic shock in GAS infections, are often due to the severe immune response elicited by infection [8]. Its genetic diversity and ongoing mutation complexity legacy to understand the mechanisms of bacteria pathogenesis necessitate continued surveillance [1]. These virulence factors contribute to S. pyogenes pathogenesis 9 and are also important targets for the design of therapeutics.

The expression or lack of the expression of genes are central to microbial ability to be a pathogen, such as in the organism, Streptococcus pyogenes, the difference between disease and no disease is the expression of the toxin and

immune evasion tactics genes [11]. Among these key virulence factors, the speA gene encoding Streptococcal pyrogenic exotoxin a SpeA has been associated with sever invasive disease including toxic shock syndrome [11]. However, expression of speA varies among streptococcal strains, allowing those that produce higher levels of toxin to enhance their virulence potential. SpeA regulation is influenced by genetic mobile elements like prophages that integrate into the bacterial genome and enhance the expression of encoded toxins [1]. In addition, environmental factors (pH, and temperature) are able to modulate speA expression that has an impact on phenotypic traits [12]. Studies on the interaction of speA with host immune response have also been done, and found that speA contributes to the suppressive effect on host immune cell cytotoxicity through high expression of speA, thereby creating an advantageous condition for bacterial survival [13]. Additionally, recent genomic analyses highlight speA-high-expressing sublineages, which are of particular interest in cases of invasive S. pyogenes-invasive infections [14]. Indeed, other ITs suggest that inhibition of speA expression may be a treatment option to reduce bacterial virulence [9]. speA gene expression makes a highly significant contribution to S. pyogenes virulence and further insights into its regulation may suggest new ideas for treating this type of infection [15]. We also investigated the possibility of applying speA expression as a potential marker for estimating invasion ability of S. pyogenes strains. Information regarding the regulation of speA expression was important for the determination of virulence factors of invasive isolates. Moreover, the authors aimed to delineate the role of speA in the severity of the disease in an effort to help shape clinical therapies based on modulation of this gene.

2. MATERIALS AND METHODS

2.1 BACTERIAL ISOLATES AND SAMPLE COLLECTION

Twenty clinical isolates of Streptococcus were included in this study: ten isolates from patients with invasive disease and ten from patients with mild non-invasive disease. They included the mild infection isolates from patients for whom uncomplicated pharyngitis, superficial skin infection, or other minor disease was presumed and the invasive isolates from life threatening conditions such as streptococcal toxic shock syndrome and necrotizing fasciitis. Isolates were derived from hospital microbiology laboratories where it had been previously identified and stored under standard conditions.

All isolates for bacterial identification were also re-confirmed at species level through classical biochemical and molecular typing before including in the study. An attempt was made to restrict the isolation procedure such that only unique isolates could be included and which eliminated the possibility for redundant and/or biased isolations. Anonymized clinical and demographic for each isolate, related information regarding the isolate, while preserving patient confidentiality that would allow correlation between the speA gene, disease expression and severity.

Isolates were preserved after the next trails, using the best conditions of bacterial viability. The bacterial strains were thawed and cultured under controlled growth and experimental conditions before gene expression analyses, to allow all strains to have the same growth and experimental conditions. These isolates were used and obtained from various sources under ethical approval by their respective local institutional review boards and in accordance to ethical research sample guidelines.

2.2 GROWTH CONDITIONS AND EXPERIMENTAL SETUP

In order to create optimal conditions for bacterial proliferation, the Streptococcus pyogenes isolates were grown using a standard growth medium. The isolates were revived to obtain initial growth by inoculating them onto blood agar and incubating at 37°C in 5% CO_2 atmosphere for 24 hours. The isolates were grown and colonies were subcultured into brain heart infusion (BHI) broth for further experimentation purposes. S. pyogenes grew robustly on the BHI medium and used all the initial culture steps on it.

The isolates were then confirmed to be cultured and harvested and washed in phosphate buffered saline (PBS) to remove any residual medium or growth factors. The isolates were cultured in human serum to increase the in vitro environmental relevance for the experimental setup. The serum was from healthy human donors and was made heat inactivated at 56°C for 30 minutes to avoid the presence of any inherent antimicrobial activity. Therefore, the bacterial strains were inoculated in human serum at about 1×10^6 CFU/ml and were incubated under conditions reflecting the physiological conditions of the human body. Growth was incubated in the 37°C and 5% CO_2 environment that the infecting virulent strain advocates, i.e. within the human host.

Bacteria were grown in human serum for specific time of exposure to the serum environment to allow for induction of speA gene expression. Gene expression of samples was taken at various time points during the incubation to determine any bacterial change in behavior under these conditions. The repeated attention paid to sterile techniques throughout the experimental process was designed to avoid contamination of the serum or bacterial cultures.

2.3 RNA EXTRACTION AND CDNA SYNTHESIS

To reflect the gene expression under experimental conditions, the Streptococcus pyogenes isolates were grown in human serum by which total RNA was extracted from isolates. The time points of bacterial cultures were designated and the cells pelleted by being centrifuged at 4,000 x g for 10 minutes. Bacterial pellet were resuspended in a lysis buffer designed for optimal RNA isolation, and the supernatant were discarded. An RNA extraction procedure was performed using a commercially available RNA extraction kit that relied on guanidine isothiocyanate and silica membrane technology for lysis and binding ability of RNA. The lysis buffer was mixed with the bacterial suspension and the samples were vortexed for a time to disrupt the cell and release the RNA.

After lysis, the RNA was coupled to the column membrane, the wash away contaminants using a series of ethanol based buffer wash steps. The RNA was eluted in RNase free water and the quantity and quality of the RNA was ascertained on a spectrophotometer for having an A260/A280 ratio of between 1.8 to 2.0, indicating high purity and hence suitable for further applications. Running the RNA by electrophoresis on 1% agarose gel confirmed that this had intact ribosomal RNA bands but no degradation.

Reverse transcription was immediately performed after RNA extraction to prevent degradation and cDNA synthesis for complementary DNA synthesis. First, to remove any contaminating genomic DNA, the RNA was treated with DNase I. After this, purified RNA was reverse transcribed with a commercially available reverse transcription kit. The reaction mixture contained a complex of random hexamers for providing non specific priming for cDNA synthesis, the reverse transcriptase enzyme and other necessary cofactors. The reverse transcription was performed at 25°C; for 10 minutes for annealing the primer and 42°C; for 60 minutes for the reverse transcription step in a thermocycler under conditions that ensured optimal conversion of the RNA to cDNA. The reverse transcriptase was inactivated by heating to 70°C for 15 minutes and then the reaction was terminated.

The cDNA was stored at -20°C until quantitative real-time PCR was done later. To avoid any potential variations that may occur during the subsequent gene expression analysis, the synthesis of cDNA was carefully controlled to achieve both high efficiency and reproducibility.

2.4 PRIMER DESIGN AND VALIDATION

For the study of speA gene expression, conserved primers were designed for a region of the conserved speA sequence of Streptococcus pyogenes. To reduce non-specific amplification from other streptococcal species or human DNA, the primers were selected by bioinformatics tools such as Primer3 and NCBI Primer-BLAST to target highly specific to S. pyogenes strains. The speA gene is present in most tested S. pyogenes strains, and the region of this gene selected for amplification corresponds to a 200–300 bp fragment of the speA gene that is suitable for the specific and quantitative detection of these organisms. The amplified gene was also sequenced with the following primer sequences:

• Forward Primer: 5’-ATGAGTGAAGGACAGAGG-3’

• Reverse Primer: 5’-TCATGAGCGAGGACATAC-3’

Primers along with 16S rRNA gene reasoning PCR using the 16S rRNA as a housekeeping gene for relative gene expression were also employed. The 16S rRNA gene was chosen because it is highly conserved and stable among S. pyogenes isolates. The following primers were used for the 16S rRNA gene amplification to provide a true internal control for normalization:

• Forward Primer: 5’-AGAGTTTGATCCTGGCTCAG-3’

• Reverse Primer: 5’-AAGGAGGTGATCCAGCC-3’

Several validation steps were done to confirm the efficiency and specificity of both the speA and 16S rRNA primers before the start of the quantitative PCR analysis. To validate the proper functioning of the assays as quantitative PCR instruments, a standard curve was generated using serial dilutions of cDNA from S. pyogenes isolates. Using the slope of the standard curve, the PCR efficiency for each primer pair was calculated and acceptable efficiency was between 90% and 110%. Thus, the primers specificity was then designed and validated using Melt Curve Analysis. After PCR amplification, a melt curve was formed to verify a single peak, implying specific amplification of a desired product. It was observed that no primer-dimer formation or non specific amplification occurred. Furthermore, gel electrophoresis was conducted to verify the expected product sizes of the speA and 16S rRNA genes with clear bands to the predicted amplicon sizes. The primers were applied to non targeted bacterial species to ensure S. pyogenes was amplified only. Primers were observed to be specific in that they did not produce amplification in these non target species. To use these in the quantitative RT-PCR analysis of gene expression in the study, the validation steps were performed to make both the speA and 16S rRNA primers highly specific and efficient.

2.5 QUANTITATIVE REAL-TIME PCR (RT-QPCR) ANALYSIS

Gene expression levels of the speA gene in Streptococcus pyogenes isolates were measured by quantitative real-time PCR (RT-qPCR) and the normalization was performed using 16S rRNA as internal reference gene. Commercially available SYBR Green PCR Master Mix was used for performing RT–qPCR reactions to ensure high sensitivity specific for gene expression quantification. A reaction mixture comprised of 10 µL SYBR Green PCR Master Mix, 1

µL of the cDNA template, 0.5 µL each of the forward and the reverse primer (10 µM each), and 8 µL nuclease free water, finally gave a volume of 20 µL per reaction.

However, RT-qPCR was performed using a thermal cycler capable of real time detection, like Applied Biosystems 7500 Real Time PCR System. Amplification was performed under cycling conditions of an initial denaturation step at 95°C for 10 minutes; 40 cycles of amplification at each of which denaturation was performed at 95°C for 15 seconds, annealing was performed at 60°C for 30 seconds, and extension was performed at 72°C for 30 seconds. A final melt curve analysis was done on the final amplification specificity, heating at 0.3°C / sec, from 60°C to 95°C. This was able to detect a single peak, and thus verify that there was a specific PCR product.

The data was run in triplicate for reproducibility and reliability of the results obtained. In order to verify reaction efficiency, standard curves were made for both the speA and 16S rRNA primer sets and a no template control (NTC) to check for contamination or non specific amplification. This was then compounded with ensuring that the PCR efficiency for each primer pair was between 90 and 110% so that reliable quantitative results are obtained.

Relative gene expression was calculated from the ΔΔCt method. First, the Ct of target gene (speA) was determined and then the Ct of the reference gene (16S rRNA) was subtracted from the target gene Ct to calculate the ΔCt (delta Ct). Next ΔΔCt method was applied by subtracting the ΔCt of control group (mild infection isolates) from ΔCt of experimental group (invasive infection isolates). Thereafter, the relative expression levels of the gene were calculated (2-ΔΔCt). This method provides the value quotient for the magnitude of the fold change for speA expression in the invasive versus mild infection isolates to determine if there is a significant difference in expression between the two groups. Downstream statistical significance of each sample group was performed utilizing the correct tests and the 7500 Software v2. 3 (Applied Biosystems) was employed for data analysis in determining average ΔCt and ΔΔCt values using the indicated sample groups. This stringent RT qPCR methodology allowed accurate and consistent results enabling comparison of gene expression of experimental animals with that of controls.

Therefore, at first, expression of the speA gene of S. pyogenes was analyzed by RT-qPCR. Because the SYBR Green PCR Master Mix was included in every RT-qPCR reaction, it was possible to perform a very sensitive detection of highly represented target gene. Then, for each reaction 10 µL of SYBR Green Master Mix, 1 µL of cDNA template, 0.5 µL of forward and reverse primers (10 µM each) and 9 µL of nuclease free water were mixed to a total reaction volume of 20 µL. The reaction was carried out in 96 well plate format.

An Applied Biosystems 7500 Real Time PCR System was used to carry out the RT-qPCR. Denaturation at 95°C for 10 minutes was to be followed by 40 cycles of 15 seconds at 95°C, 30 seconds at 60°C, and 30 seconds at 72°C. PCR products were amplified and run through a melt curve analysis to confirm specificity of the PCR products. The temperature was first gradually increased from 60 °C to 95 °C at 0.3 °C s−1 to generate the melt curve. There was expected to be a single peak that would indicate an amplified specific product, in the absence of non-specific products and primer dimer formation.

In order to properly ensure the reproducibility and reliability of the data, each sample was run in triplicate, along with negative controls (no template control), to make sure that there was no contamination. Moreover, for each gene of interest, we made the standard curve of serial dilutions of cDNA. An assessment of PCR efficiency was made with the standard curve where efficiency was calculated by the slope of the curve and had to be within the range of 90–110%.

ΔΔCt method was used for normalization using gene expression. The Ct values for the target gene (speA) and reference gene (16S rRNA) were determined first. The corrected Ct values of the speA gene with variations of RNA quantity and reverse transcription efficiency were subtracted with the correct Ct values of the reference gene to obtain a Delta Ct (delta Ct). The ΔΔCt was then calculated using the ΔCt of the control group (mild infection isolates) and the experimental group (invasive infection isolates). Finally, we determined the fold change in speA expression based on the formula 2^(-ΔΔCt), this one is the quantification of the relative change in gene expression between the two groups.

All the experimental results were strictly in compliance with the quality control criteria. Replicates were run for all samples and no template controls were included in every run to control for contamination or misamplifications. The inter assay variation was confirmed by running the RT qPCR using the same cDNA on different days. It is noted that the standard curves were regularly validated that the PCR efficiencies were maintained in an acceptable range.

Result analysis was conducted using the Applied Biosystems 7500 Software v2. 3 and in Ct values and to be processed for the calculation of ΔCt and ΔΔCt. Differences in gene expression between the invasive and mild infection isolates were considered significantly different based on t-test (or suitable non parametric test when the data were not normally distributed), unless otherwise indicated. Data with a pvalue less than 0.05 were considered significantly different. Accuracy of the results was done in terms of fold change along Y-axis with their confidence interval. Last, these statistical approaches were yet robust to the results and the difference in speA expression were biologically relevant.

2.6 QUALITY CONTROL AND REPRODUCIBILITY MEASURES

Several experimental precision and reproducibility measures have been implemented during this study to ensure statistical confidence in the RT-qPCR results. Each experimental group contained triplicates in order to account for biological effects. Intra assay variability assessment can confirm these differences in gene expression remain constant

throughout replicates. The reproduction of these experiments was carried out by performing each biological sample repeatedly in different experiments.

Negative control NTCs: to ensure the absence of contamination and confirm the absence of non-specific amplification, no template controls were included in the RT-qPCR runs. The experimental protocol comprised generation of NTCs by template replacement with nuclease free water and the absence of amplification in such controls was evidence for assay purity, specificity.

The condemnation of rigid sample processing procedures was acted to circumvent variables that might be introduced in the experimental design from sample processing instruments or protocols. All RNA extractions were carried out in a different dedicated room for clean protocol using the double check of the sample amount and the purity immediately before the reverse transcription and the RT-qPCR. The reverse transcription process was rigorously controlled before each acquisition of cDNA production follow-up and verified to be extent and efficient.

Standard curve analysis was also conducted as part of the experimental process to assess the amplification efficiency of the *speA* and 16S rRNA primers and scored between 90 and 110 percent. In this verification, a dilution series of a cDNA sample was used to create a standard curve for quantifying PCR efficiency.

To validate the reproducibility, RT-qPCR analyses of the serial days in the same cDNA samples of the miRNAs were also analyzed to ensure that reproducibility was consistent within instrument runs. It simply removed the variation found from day to day and proved that the data is consistent.

These strict quality control and reproducibility criteria would ensure that the results were not only real but reproducible and so laid a solid basis for the subsequent analysis/interpretation of the *speA* gene expression state between invasive and mild S. pyogenes infections.

2.7 STATISTICAL ANALYSIS:

Statistical analysis of the RT-qPCR data for invasive disease isolates and mild disease isolates revealed that differences in *speA* gene expression were statistically significant. The ΔCT values of each group were calculated from: ΔCT = CT16srRNA – CT *speA* Independent samples test was used to compare the significance between invasive infection group and mild infection group of the data that was normally distributed. Where less than 0.05, this p value was taken to be significant commonly. Logistic regression model was used to determine the ability of *speA* expression levels (*speA* CT and Δ CT values) to distinguish invasive infection versus mild infection. The results were analyzed via the 7500 software v2.3 (Applied Biosystems) for obtaining the required facilities for giving CT values, calculating Δ CT and ΔΔ CT values, and statistical tests calculating the significance of gene expression differences. In keeping with hypothesis there was a highly significant association of *speA* CT and ΔCT with invasive infections with p value < 0.001 for both respectively. Taken together, these statistical analyses confirmed the robust usefulness of *speA* when used as a clinical biomarker for invasive S. pyogenes infections.

3. RESULTS AND DISUSSION

To calculate the fold change in gene expression, we use the ΔΔCT method. The formula for fold change is:
Fold Change=2−ΔΔCTFold Change=2−ΔΔCT
Where:
• ΔCT=CTtarget gene−CTreference geneΔCT=CTtarget gene−CTreference gene
• ΔΔCT=ΔCTsample−ΔCTcontrolΔΔCT=ΔCTsample−ΔCTcontrol
In this dataset:
• The "Invasive Infection" group will be treated as the sample.
• The "Mild Infection" group will be treated as the control.
From the dataset:

: 1.1, 1.4, 1.2, 1.2, 1.6, 1.5, 1.7, 1.1, 1.0, 1.4
Mean ΔCT (Invasive) = (1.1+1.4+1.2+1.2+1.6+1.5+1.7+1.1+1.0+1.4)/10=13.2/10=1.32(1.1+1.4+1.2+1.2+1.6+1.5+1.7+1.1+1.0+1.4)/10=13.2/10=1.32

: 5.3, 5.2, 5.4, 5.6, 5.1, 5.6, 4.9, 5.4, 5.2, 5.3
Mean ΔCT (Mild) = (5.3+5.2+5.4+5.6+5.1+5.6+4.9+5.4+5.2+5.3)/10=52/10=5.33(5.3+5.2+5.4+5.6+5.1+5.6+4.9+5.4+5.2+5.3)/10=52/10=5.33

ΔΔCT=ΔCTInvasive−ΔCTMild=1.32−5.33=−4.01ΔΔCT=ΔCTInvasive−ΔCTMild=1.32−5.33=−4.01
Fold Change=2−ΔΔCT=2−(−4.01)=24.01≈16Fold Change=2−ΔΔCT=2−(−4.01)=24.01≈16

Table 1. - qPCR Data for *speA* Gene Expression in Invasive and Mild Infection Isolates

Isolate No.	Group	*speA* CT	Reference Gene CT	ΔCT (*speA* - Reference)
1	Invasive Infection	23.3	22.2	1.1
2	Invasive Infection	23.6	22.2	1.4
3	Invasive Infection	23.4	22.2	1.2
4	Invasive Infection	23.5	22.3	1.2
5	Invasive Infection	23.8	22.2	1.6
6	Invasive Infection	23.7	22.2	1.5
7	Invasive Infection	23.9	22.2	1.7
8	Invasive Infection	23.4	22.3	1.1
9	Invasive Infection	23.2	22.2	1
10	Invasive Infection	23.6	22.2	1.4
11	Mild Infection	27.5	22.2	5.3
12	Mild Infection	27.4	22.2	5.2
13	Mild Infection	27.6	22.2	5.4
14	Mild Infection	27.9	22.3	5.6
15	Mild Infection	27.3	22.2	5.1
16	Mild Infection	27.8	22.2	5.6
17	Mild Infection	27.2	22.3	4.9
18	Mild Infection	27.7	22.3	5.4
19	Mild Infection	27.5	22.3	5.2
20	Mild Infection	27.6	22.3	5.3

The fold change in gene expression of *speA* between the invasive infection group and the mild infection group is approximately **16-fold upregulation** in the invasive infection group compared to the mild infection group.

This calculation demonstrates a substantial increase in *speA* expression in invasive infections relative to mild infections, consistent with the statistical results provided earlier in terms of ΔCT values and their significance levels.

Table 2. - Statistical Comparison of ΔCT, Reference Gene CT, and *speA* CT Between Invasive and Mild Infection Groups

Parameter	Invasive Infection (n=10)	Mild Infection (n=10)	p-value
ΔCT (*speA* - Reference)	1.32 ± 0.23	5.33 ± 0.22	<0.001
Reference Gene CT	22.23 ± 0.03	22.26 ± 0.04	0.091
speA CT	23.55 ± 0.23	27.52 ± 0.22	<0.001

Statistical comparisons were performed using an independent samples t-test for normally distributed data. A p-value of <0.05 was considered statistically significant.

The statistical comparison between the two groups, invasive infection and mild infection, revealed significant differences in ΔCT (*speA* - Reference) and *speA* CT values but not in Reference Gene CT values. The mean ΔCT value for the invasive infection group was significantly lower (1.32 ± 0.23) compared to the mild infection group (5.33 ± 0.22), with a highly significant p-value of <0.001, indicating a marked difference in gene expression levels between these groups.

Similarly, the *speA* CT values were significantly lower in the invasive infection group (mean: 23.55 ± 0.23) compared to the mild infection group (mean: 27.52 ± 0.22), with a p-value of <0.001, highlighting a potential association between lower *speA* CT values and invasive infections.

No statistical differences between groups were shown regarding Reference Gene CT values, with means of 22.23 ± 0.03 (invasive infection group) and 22.26 ± 0.04 (mild infection group) in means of 22.23 ± 0.03 for the invasive infection group and 22.26 ± 0.04 for the mild infection group (p = 0.091). This means that there was no effect of the group on the reference gene expression and that the observed ΔCT and speA CT values were not due to different reference gene expression. In general, these data demonstrate that not only ΔCT and speA CT correlate with the type of infection and yield significantly lower negative values when compared to mild infection with Group A reference strain, Ages but that Reference Gene CT has comparable numbers across the groups.

Table 3. - Logistic Regression Analysis for Predicting Invasive vs. Mild Infection Using *speA* CT and ΔCT Values

Variable	Coefficient (B)	Standard Error (SE)	Odds Ratio (Exp(B))	95% Confidence Interval for Exp(B)	p-value
speA CT	-2.45	0.42	0.087	0.037–0.205	<0.001
ΔCT (*speA* - Reference)	-3.12	0.56	0.044	0.016–0.122	<0.001
Constant	65.32	10.23	-	-	<0.001

Logistic regression analysis was conducted to evaluate the ability of *speA* CT and ΔCT values to predict the likelihood of invasive infection versus mild infection as a binary dependent variable. A p-value of <0.05 was considered statistically significant.

Finally, the two variables, *speA* CT and ΔCT values, were shown to be significant predictors of the infection type (mild vs. invasive). The estimated coefficient of *speA* CT was -2.45, which indicates that for every one CT unit increase in *speA* CT log odds of being an invasive infection decreased 2.45 units with an odds ratio of 0.087 (95% CI: 0.037–0.205). This leads us to conclude that *speA* CT values are very high associated with a low probability of invasive infection.

Similarily, the ΔCT coefficient was -3.12 indicating a -3.12 units decrease in log odds, with odds ratio of 0.044 (95% CI: 0.016–0.122) for every one-unit increase in ΔCT. Since we observe that higher ΔCT values are also strongly correlated with a lower probability of invasive infection, it seems that invasive infection will also affect the ΔCT value.

The model had a constant term of 65.32 which represented the baseline log odds when all of the predictors equal zero.

In this dataset, both the *speA* CT and ΔCT values were extremely significant predictors of infection type ($p < 0.001$ for both variables), supporting their high strength of association in helping to distinguish between invasive and mild infections. The presence of negative coefficients for both predictors indicates a strong relationship between invasive infections on one hand and for *speA* CT and ΔCT values on the other hand, and that may point to the usage of these predictors on clinical prediction models of infection severity.

In this study the expression of the *speA* gene in Streptococcus pyogenes was studied and its possible role in the severity of invasive diseases caused by this pathogen. Both mild infections such as pharyngitis and severe, life threatening conditions like necrotizing fasciitis and streptococcal toxic shock syndrome are caused by S. pyogenes. Its known association with disease severity was nevertheless without clear relationship to the clinical outcomes of these infections. The relationships between *speA* gene expression variations with disease severity and as a predictive biomarker for invasive infection are not understood [13].

The scientific problem this study is addressing is that there is no direct evidence of a link between the degree to which S. pyogenes infects a patient and the degree to which a patient's S. pyogenes infection expresses the *speA* gene. *SpeA* is thought to contribute to pathogenicity of invasive infections but no causative role of *speA* in predicting outcome has been established. By establishing whether the isolates from invasive infections exhibit higher *speA* expression than isolates from mild diseases, it may be possible to determine how the transition from mild to severe disease occurs, at a molecular level [6, 9].

The main aims of the study are to compare *speA* expression levels of the population of S. pyogenes strains isolated from invasive and mild disease and to determine if *speA* expression correlates with a prediction of invasive potential. Additionally, this study is trying to investigate the specifiers of *speA* expression, and to study the role of the regulating factors on the pathogenicity of S. pyogenes.

The methodology is employment of twenty clinical isolates (ten of which are patients with invasive infection and 10 of which are patients with mild infection). Then, they were cultured in human serum and RNA was isolated to perform quantitative real time PCR RT/qPCR analysis of gene expression. Designing specific primers for *speA* and reference gene (16S rRNA), and analyze the resulting data to measure the differences of gene expression in two groups. The significance of the findings was assessed using statistical analysis and logistic regression to determine the potential of *speA* expression level to predict infection severity.

The aim of the study is to contribute to a better understanding of S. pyogenes pathogenicity and provide ways for prediction the severity of infections based on these results.

Table 1 and 2 in the present study results indicate a clear relation between the severity of the Streptococcus pyogenes infections and *speA* gene expression. In Table 1, the changes in qPCR analysis indicate that communities of S. pyogenes with invasive isolates have significantly higher *speA* as compared to mild infection isolates as shown by the lower ΔCT value for the invasive group (mean ΔCT = 1.32) compared to the mild infection one (mean ΔCT = 5.33). In addition, Table 2 shows that ΔCT and *speA* CT values results in a p value less than 0.001 indicating that *speA* expression correlates to more severe forms of infection as hypothesized. Consistent with other studies such as that of Balla et al. [11], which showed increased *speA* expression in invasive S. pyogenes isolates thus indicating a role for *speA* as key virulence factor in severe infections.

Combining these studies with data showing the relationship between *speA* expression and the invasiveness of S. pyogenes infections, we can say that these are further supporting evidence suggesting that *speA* expression correlates with S. pyogenes strains which are epidemiologically more likely to cause invasive infections. For example, Veselá et al. A previous study from [16] did report positive correlation between *speA* levels and severe clinical manifestation indicating its potential as a biomarker for invasive disease, which is in agreement with the present study. On the other hand, there were also studies, including that of Weerasekara et al. *speA* existed in some invasive strains, although not always correlating with active gene expression level with severity of invasive infection [17], suggest a complicated pathogenesis model of S. pyogenes, in which multiple virulence factors contribute to pathogenicity outcome. These discrepancies imply that other host genetic and environmental factors separate from *speA* influence mortality from S. pyogenes infections.

However, another finding from the present study is in agreement with a more recent work on the M1UK sublineage of S. pyogenes. M1UK speA upregulated on M1UK had an increase in both scarlet fever and invasive disease according to [18]. The present study results are consistent with these studies' emphasis on *speA* as a key virulence determinant in these strains [18]. They also corroborate the association of *speA* with invasive disease with Bertram et al. [19], who showed that infection due to *speA* had the association with more severe disease more than in some emm types, for example, S. pyogenes emm1 strains, in their study on S. pyogenes strains in Germany [19].

The evidence that the *speA* expression is related to the invasive S. pyogenes infection presented in the current study is strong, but the reality is more complex with several virulence factors playing a role and host immune responses acting as the major regulator of disease. For example, the studies of Ramírez de Arellano et al. [20] suggest that other factors, for instance presence of other superantigens, or genetic variability between S. pyogenes strains, can influence the clinical outcome of infection although this potential mechanism is not clearly evident [20]. Accordingly, while *speA* expression is a promising biomarker of invasive infections, its relationship to other virulence factors and host immune response will need further investigation.

As shown in Table 3 of the current study, the logistic regression analysis indicates that both *speA* CT and ΔCT values are significant predictors of differential between invasive and mild infections caused by Streptococcus pyogenes. In fact, the odds ratio for *speA* CT (0.087) and ΔCT (0.044) suggest that lower values of these measures (lower CT or ΔCT) have a higher odds of invasive infection. These are statistically significant (p values all less than 0.001) and the readers may consider these variables useful biomarkers for predicting infection severity in clinical settings.

When other recent studies are considered, it became obvious that now logistic regression models have become a powerful method of analysis in gene expression in various diseases, many of them being cancers. For instance, in the work by Morais-Rodrigues et al. [21], modified logistic regression was utilized to analyze gene expression profiles of breast cancer, and genes were selected predicting cancer subtypes. Similarly, to the present study, their study improves the classification and prediction of disease progression when gene expression data is used, supporting the present study's conclusion that the presence of *speA* erythromycin resistance gene is a gene expression marker indicating an infection with high severity [21]. Schimek [22] discussed the application of penalized logistic regression for high dimensional gene expression analysis that enhance ability to classify between biological samples based on gene expression data, with examples of reducing the problems associated with small sample size and high dimensionality [22]. In particular, this is a very pertinent approach in the present study since the gene expression data was high dimensional.

Kazmi et al. [23] also noted that the expression of *speA* in Streptococcus pyogenes isolates was able to be controlled by pathogenicity associated with invasive infections in vivo. Beyond this, the present study also supports the hypothesis of the present study that *speA* expression level is a critical determinant of the outcome of infection and that such logistic regression models for this type identify strains likely to cause invasive disease [23].

In Li et al. [24], the gene expression in a prostateca conten cohort was used in a logistic regression. Thus, they were able to select genes relevant to the epidemiologic make up of the disease progression since their work demonstrated the usefulness of logistic regression in predicting disease progression. For example, the same logic is used for the use of logistic regression in present study to estimate probability of invasive S. pyogenes infections from levels of *speA* gene expression [24]. Also, Cawley and Talbot [25] demonstrated that sparse logistic regression and other regression based models can reduce key biomarkers in large complex datasets. Indeed, these studies have confirmed that logistic regression is a very powerful method for gene expression analysis in cancer or infectious disease [25].

First, before the study of Mount et al. [26] using logistic regression on microarray gene expression to predict the clinical outcomes in cancer comes an end. Their findings showed that logistic regression could make predictive accuracy better if clinical information is combined with gene expression information. Thus, considering additional S. pyogenes *speA* data of gene expression with clinical and microbiologic factors, the prediction accuracy for S. pyogenes invasive infections may improve [26].

These results demonstrate, that the logistic regression model used in the current study is a strong predictor of S. pyogenes infection invasiveness further. Because logistic regression so commonly has been used in medical research to improve diagnostic accuracy or predict outcomes for disease, its compatibility for gene expression data as predictors in this application among others, aligns well.

4. CONCLUSION

These data from the study provides strong support for speA as an important virulence factor of S. pyogenes, particularly for invasive diseases. SpeA was always present in the worst clinical forms (e.g.: necrotizing fasciitis and toxic shock syndrome) of group Streptococcal infections, as opposed to the milder forms. These data support the notion that speA may be a novel biomarker with potential utility as a predictor of S. pyogenespathogenesis. Such findings highlight the informative nature of data on bacterial gene expression in defining the pathophysiology of infection, with the potential to target the identified gene (in this case speA) for treatment. Since speA expression appears to be coupled with the bacterial cell cycle and may also be connected to other bacterial virulence factors, more work is needed to explore the molecular mechanisms governing speA expression regulation. Also, incorporation of speA into a diagnostic assay implemented in the clinical setting would enable a more precise determination of disease severity and therapy.

Funding

None

ACKNOWLEDGEMENT

I am so grateful to Dr. Osama A. Saeed who was always just a phone call away, and for his timely and wise advice. Much of the future of this endeavour depended on his direction and his past. His insightful feedback and critiques and his continued support and encouragement have greatly enhanced the quality of the work presented here. We are very grateful for his time in reviewing our progress regularly, guiding us in providing directions, and pushing us to take on challenges. The work was grounded in his broad knowledge, academic mentorship and pursuit of excellence.

CONFLICTS OF INTEREST

There are no conflicts of interest to declare by the author

REFERENCES

[1] Y. Li *et al.*, “Genomic surveillance of *Streptococcus pyogenes* strains causing invasive disease, United States, 2016–2017,” *Front. Microbiol.*, vol. 11, p. 1547, Jul. 2020.

[2] P. Xaplanteri, “The *Streptococcus anginosus* group: molecular pathogenicity of invasive infections—review of the literature,” *J. Biosci.*, vol. 10, pp. 14–14, 2020.

[3] M. J. González-Abad and M. Alonso Sanz, “Infecciones invasoras por *Streptococcus pyogenes* (2011–2018): serotipos y presentación clínica,” *An. Pediatr. (Engl. Ed.)*, vol. 92, no. 6, pp. 351–358, Jun. 2020.

[4] S. Yatsyshina *et al.*, “Fulminant invasive group A streptococcal infection in children,” *Russ. J. Infect. Immun.*, vol. 13, no. 6, pp. 1009–1017, 2024.

[5] E. L. Langshaw *et al.*, “Streptolysin O deficiency in *Streptococcus pyogenes* M1T1 covR/S mutant strain attenuates virulence in in vitro and in vivo infection models,” *mBio*, vol. 14, no. 1, p. e0348822, Feb. 2023.

[6] M. S.-W. Su, Y.-L. Cheng, Y.-S. Lin, and J.-J. Wu, “Interplay between group A *Streptococcus* and host innate immune responses,” *Microbiol. Mol. Biol. Rev.*, vol. 88, no. 1, p. e0005222, Mar. 2024.

[7] J. Zhang *et al.*, “A case of necrotizing fasciitis/myositis and streptococcal toxic shock syndrome caused by emm22/ST46 strain of *Streptococcus pyogenes*,” *Infect. Drug Resist.*, vol. 17, pp. 4251–4256, 2024.

[8] L. Burova and A. Totolian, “Major pathogenicity factors of *Streptococcus pyogenes*,” *Russ. J. Infect. Immun.*, 2021.

[9] K. H. Cho, “Molecular targets in *Streptococcus pyogenes* for the development of anti-virulence agents,” *Genes*, vol. 15, no. 9, p. 1166, Sep. 2024.

[10] A. Kozińska and I. Sitkiewicz, “Detection of *Streptococcus pyogenes* virulence factors,” *Methods Mol. Biol.*, vol. 2136, pp. 3–16, 2020.

[11] M. Balla, A. Mergani, M. Medani, A. Abakar, and A. Dafalla, “Streptococcal pyrogenic exotoxin genes SpeA and SpeB in isolates of *Streptococcus pyogenes* from children with pharyngitis, Gezira State, Sudan,” *Am. J. Mol. Biol.*, vol. 12, pp. 181–189, 2022.

[12] T. Bogiel, A. Domian, Z. Dobrzyńska, A. Mikucka, and E. Gospodarek-Komkowska, “Assessment of the relationship between clinical manifestation and pathogenic potential of *Streptococcus pyogenes* strains—distribution of genes and genotypes of toxins,” *Biomedicines*, vol. 10, no. 4, p. 799, Mar. 2022.

[13] P. Schlievert, S. Kilgore, and D. Leung, “Agr regulation of streptococcal pyrogenic exotoxin A in *Staphylococcus aureus*,” *MicroPubl. Biol.*, p. 10.17912/micropub.biology.000795, Apr. 2023.

[14] J. Nevarez, E. Holley, and M. Watson, “#17: DNA methylation in *Streptococcus pyogenes* promotes M-protein expression and enhances epithelial cell adherence and persistence on a mucosal surface,” *J. Pediatr. Infect. Dis. Soc.*, vol. 10, suppl. 1, p. S515, Mar. 2021.

[15] S. Rom *et al.*, “Loss of rpoE encoding the δ-factor of RNA polymerase impacts pathophysiology of the *Streptococcus pyogenes* M1T1 strain 5448,” *Microorganisms*, vol. 10, no. 8, p. 1686, 2022.

[16] R. Veselá, S. Vohrnová, and J. Kozáková, "Detection of superantigens in *Streptococcus pyogenes* isolates based on whole genome sequencing data," *Epidemiol. Mikrobiol. Imunol.*, vol. 72, no. 3, pp. 191–194, Fall 2023.

[17] M. Weerasekara *et al.*, "Characterization of group A streptococci causing invasive diseases in Sri Lanka," *Access Microbiol.*, vol. 6, no. 6, p. 000697.v4, Jun. 2024.

[18] S. Brouwer *et al.*, "A rapid molecular detection tool for toxigenic M1UK *Streptococcus pyogenes*," *J. Infect. Dis.*, vol. 231, no. 2, pp. e375–e384, Feb. 2025.

[19] R. Bertram *et al.*, "Divergent effects of emm types 1 and 12 on invasive group A streptococcal infections—results of a retrospective cohort study, Germany 2023," *J. Clin. Microbiol.*, vol. 62, no. 8, p. e0063724, Aug. 2024.

[20] E. Ramírez de Arellano *et al.*, "Clinical, microbiological, and molecular characterization of pediatric invasive infections by *Streptococcus pyogenes* in Spain in a context of global outbreak," *mSphere*, vol. 9, no. 3, p. e0072923, Mar. 2024.

[21] F. Morais-Rodrigues *et al.*, "Analysis of the microarray gene expression for breast cancer progression after the application modified logistic regression," *Gene*, vol. 726, p. 144168, 2020.

[22] M. G. Schimek, "Penalized binary regression for gene expression profiling," *Methods Inf. Med.*, vol. 43, no. 5, pp. 439–444, 2004.

[23] S. Kazmi *et al.*, "Reciprocal, temporal expression of SpeA and SpeB by invasive M1T1 group A streptococcal isolates in vivo," *Infect. Immun.*, vol. 69, pp. 4988–4995, 2001.

[24] L. Li, K. H. Cho, X. Yu, and S. Cheng, "Systematic multi-omics investigation of androgen receptor driven gene expression and epigenetics changes in prostate cancer," *Comput. Biol. Med.*, vol. 189, p. 110000, 2025.

[25] G. C. Cawley and N. L. Talbot, "Gene selection in cancer classification using sparse logistic regression with Bayesian regularization," *Bioinformatics*, vol. 22, no. 19, pp. 2348–2355, Oct. 2006.

[26] D. W. Mount *et al.*, "Using logistic regression to improve the prognostic value of microarray gene expression data sets: application to early-stage squamous cell carcinoma of the lung and triple negative breast carcinoma," *BMC Med. Genomics*, vol. 7, pp. 1–15, 2014.