INTRODUCTION

To diagnose latent tuberculosis infection (LTBI), the tuberculin skin test (TST) has traditionally been used. However, TST has low specificity, because it could be influenced by Bacillus Calmette–Guérin vaccination and non-tuberculous mycobacterial infection. Interferon gamma release assay (IGRA) is an alternative to TST for diagnosing LTBI that is not influenced by these factors [1–4].

There are two kinds of commercial interferon gamma release assay, QuantiFERON–TB Gold In-Tube (QFT)
Assay (Cellestis/Qiagen, Carnegie, Australia) and the T-SPOT.TB Assay (Oxford Immunotec, Abingdon, UK). The QFT assay is an enzyme linked immunosorbent assay (ELISA)-based test that detects interferon gamma released by activated T lymphocytes with peptides encoded by the region of deletion 1 (RD1) gene, the 6-kDa Mycobacterium tuberculosis protein early secreted antigenic target 6 (ESAT-6), and the 10-kDa culture filtrate protein (CFP-10), as well as peptides from one additional antigen (TB7.7 [Rv2654c], which is not an RD1 antigen), in an in-tube format. An individual is considered positive for M. tuberculosis infection if the interferon gamma (IFN-\(\gamma\)) response to tuberculosis (TB) antigens is above the test cut-off after subtracting the background IFN-\(\gamma\) response of the negative control. The T-SPOT.TB assay is an enzyme-linked immunosorbent spot assay performed on separated and counted peripheral blood mononuclear cells (PBMCs). The result is reported as the number of IFN-\(\gamma\)-producing T cells (spot-forming cells). The T-SPOT.TB assay is more laborious, because it requires manual processing of PBMCs separation and normalizing the number of PBMCs in each test well; additionally, the counting needs to be interpreted with the naked eye (5-8).

The QFT is a qualitative determination test based on quantitative results of the released IFN-\(\gamma\) levels. This is generally used as an IGRA in Korea. As stated in 'The Action Strategy for TB-Free Korea' last March, almost two million Koreans are to be screened for LTBI, exclusively with QFT, because it enables large-scale screening with its automation and analytical easiness. There are two types of automated ELISA analysers for QFT, DS2 (Dynex Technologies, Chantilly, VA, USA) and Agility (Dynex Technologies).

![Fig. 1. Overview of the QuantiFERON-TB Gold In-Tube (Cellestis/Qiagen, Carnegie, Australia) assay, performed using (A) manual ELISA method, (B) DS2, and (C) Agility. Abbreviation: ELISA, enzyme linked immunosorbent assay.](image)
Agility was released as a system with a higher throughput than that of DS2, which allows the analysis of only two 96-well plates per batch (Fig. 1). Agility has been used in many foreign laboratories, and last year, it was introduced for the first time in Korea. However, little data exists on the performance of Agility; thus, the aim of the study was to evaluate the accuracy and speed of Agility for IGRA (Fig. 2).

MATERIALS AND METHODS

1. Materials

Pooled plasma was prepared using positive and negative samples obtained after IGRA testing with Agility, separately. Nil plasma samples obtained after testing were pooled and used for sample dilution. The QFT assay kit was used for all tests.

2. Methods

1) Device familiarization period

Prior to precision evaluation, the first five operating days were spent familiarizing with the device, including aspects, such as the set-up, operation, and maintenance during routine laboratory use.

2) Quality control

Quality control (QC) was performed for every batch of samples.
with positive and negative QC materials provided by the manufacturer [9].

3) Standard curve generation

For every batch of tested samples, a standard curve was generated and confirmed by duplicating it with the kit standard reagent, covering the range from 4 to 0 IU/mL, according to the manufacturer’s protocol (STANDARD 1:4 IU/mL, STANDARD 2:1 IU/mL, STANDARD 3:0.25 IU/mL, STANDARD 4:0 IU/mL) [9]. Regression analysis was performed with the Microsoft Office Excel package (Microsoft Corp., Redmond, WA, USA). In each batch, the standard curve was checked for compliance with the manufacturer’s guideline.

4) Precision evaluation

For precision evaluation, repeatability (formerly termed within-run precision), between-run precision, and between-day precision were analysed on three different Agility instruments at the same time and with the same samples. Two types, each, of high- and low-concentration samples were serially tested in duplicates via two separated runs at 4-hour-time intervals repeatedly for 20 days. The standard deviation (SD) and coefficient of variation (CV) of all data observed were calculated and analysed according to the Clinical and Laboratory Standards Institute (CLSI) EP5-A3 guideline [10].

5) Linearity

The method for measuring IFN-\(\gamma\) concentration by the QFT ELISA technique has been demonstrated to be linear from 0 to 10 IU/mL [11]. The linearity study was performed by placing five replicates of eleven plasma pools of known IFN-\(\gamma\) concentrations randomly on the ELISA plate [11]. Linearity was analysed according to the CLSI EP6-A guideline [12]. The pooled plasma sample with a high IFN-\(\gamma\) concentration of up to 10.837 IU/mL, identified in the quantitative results via repeated tests, was prepared and diluted at regular concentration intervals using the pooled nil plasma. Each sample was tested in triplicate and the average values were used for recovery analysis and for comparing the expected and measured concentration values. The SD and CV were also calculated by measuring the concentration of the test materials. Linear regression was calculated from the average and represented as \(y=ax+b\) (\(y\): determined concentration; \(x\): expected concentration).

6) Comparison with DS2 and manual method

Results of Agility were compared to the conventional ones of DS2 or the manual test methods, according to CLSI EP12-A2 and CLSI EP9-A3 guidelines [13,14]. A total of 116 samples, in total, were evaluated and duplicated, and the concordance rate was shown as a percentage.

7) Cut-off evaluation

The cut-off of the QFT test was evaluated in the laboratory according to the CLSI EP12-A2 guideline to identify the range of uncertainty near the cut-off values as the \(C_{50}-C_{95}\) interval [13]. Based on the concentration of pooled plasma, the specimen corresponding to the cut-off value of 0.35 IU/mL was prepared and its compatibility with the \(C_{50}\), that would theoretically yield 50% positive results and 50% negative results when 40 replicates of a single sample are tested, was validated.

After the cut-off was identified to be compatible with the \(C_{50}\) value, \(C_{50} \pm 20\%\) samples were prepared and tested in triplicates by repeating the tests 40 and 100 times each. For the uncertainty test, two samples corresponding to the concentration of \(C_{50} \pm 10\%\) were prepared and the verification was performed in triplicates by 40 repeated tests.

8) Turnaround time and throughput

Turnaround time for a batch test was investigated as the average time taken for IGRA from submission of the centrifuged plasma sample to completion, i.e., reporting of results, excluding pre-analysis processes, like blood sampling and 20-hour antigen activation. Throughput performance per batch (29 tests each for two plates per batch) and per day were measured and calculated for a routine working time of 8 hours, and were compared with the estimated performances of Agility and DS2.
9) Statistical analysis

Results for precision, linearity, correlation with other methods, and cut-off evaluation were analysed according to the CLSI EP5-A3, EP6-A, EP9-A3, and EP12-A2 guidelines. Statistical analysis was performed using Microsoft Office Excel (Microsoft Corp.).

RESULTS

1. Standard curve generation

Using the standard reagent from the QFT kit, a valid standard curve was generated and duplicated for each batch. The mean optical density (OD) value for STANDARD 1 was above 0.600 (1.253, SD=0.042). The %CV for STANDARD 1 to STANDARD 3 was to be calculated within 10%. Replicate OD values for STANDARD 3 and STANDARD 4 did not vary from their mean by more than 0.040 OD units.

2. Precision

The CV ranged from 2.92% to 4.0% and 2.48% to 3.42% in the repeatability analysis, 7.79% to 11.17% and 7.01% to 10.84% in the between-run precision tests, and 11.93% to 14.84% and 9.69% to 12.83% in the between-day precision tests, at low and high concentration levels, respectively (Table 1).

3. Linearity

In serial number (S/N) 133, linearity was sustained from 0 to 10.541 and the regression equation was \( y=1.0014x+0.0315 \). In S/N 134, linearity was sustained from 0 to 10.541 and the regression equation was \( y=0.9841x–0.0364 \). In S/N 137, linearity was sustained from 0 to 10.541. The coefficient of determination (R²) was above 0.99 for all three instruments (Fig. 3).

4. Comparison analysis with DS2 and manual methods

Comparison analysis was replicated 2 or 3 times with negative and positive samples using the DS2 and manual
methods. The concordance rate of the three Agility instruments with the conventional methods was 99.99% (Table 2). There was only one discordant case of low INF-γ levels near the cut-off value (Tables 2, 3).

5. Cut-off evaluation
The sample, with its concentration level adjusted to 0.35 IU/mL, showed consistent positive results within the range of 20-26 times among the total of 40 repetitions; these results were compatible with the proper C₅₀ value. The C₅₀−20% and C₅₀+20% samples showed all negative and all positive results, respectively, in the triplicated 40-test sets, indicating that C₅₀±20% contains the C₅-C₉₅ interval, with 85.8% confidence. This was also repeatedly demonstrated with the 100-test sets. The C₅₀−20% and C₅₀+20% samples showed negative and positive results, respectively, 96-100 times in the triplicated 100-test sets. On the other hand, the C₅₀−10% and C₅₀+10% samples showed inconsistent negative and positive results, respectively, in the range of 35-40 times in the triplicated 40-test sets, suggesting that the C₅-C₉₅ interval was not contained in the C₅₀±10% range (Table 4).

6. Turnaround time and throughput
The turnaround time was 3.84 hours from the submission of samples to the reporting of results for a batch test in Agility. The estimated throughput values of DS2 and Agility were 116 tests and 464 tests, respectively, within an 8-hour period during regular working hours; this suggests that Agility is expected to perform 4 times the number of tests that DS2 can perform. However, the actually measured throughput of Agility was 290 tests

<p>| Table 2. Comparison analysis of Agility with DS2 and manual methods |
|------------------------|----------------|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>Serial no. of instrument</th>
<th>No. of test</th>
<th>No. of concordant cases</th>
<th>No. discordant cases</th>
<th>Overall concordance rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>133</td>
<td>453</td>
<td>453</td>
<td>0</td>
<td>100.0</td>
</tr>
<tr>
<td>134</td>
<td>347</td>
<td>346</td>
<td>1</td>
<td>99.7</td>
</tr>
<tr>
<td>137</td>
<td>232</td>
<td>232</td>
<td>0</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>1,032</td>
<td>1,031</td>
<td>1</td>
<td>99.99</td>
</tr>
</tbody>
</table>

Fig. 3. Linearity results of Agility. The linearity regression was sustained from 0 to 10.541 on the three Agility instruments. $R^2$ value presented above 0.99. Abbreviations: S/N, serial number; IFN, interferon gamma.

<table>
<thead>
<tr>
<th>Table 3. Interferon gamma release assay result of the discordant case in the comparison analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>134-1</td>
</tr>
<tr>
<td>134-1</td>
</tr>
<tr>
<td>134-1</td>
</tr>
<tr>
<td>134-1</td>
</tr>
</tbody>
</table>

Abbreviation: TB Ag, tuberculosis antigen.
for a routine working time, which is 2.5 times more than that for DS2 (Table 5). Sequential insertion of new plates for testing is available every 50 minutes (approximately) after testing for a previous plate is completed; however, there were unexpected, frequent time lags observed while finishing tests, thereby reducing the overall throughput to a value less than expected (Table 6).

**DISCUSSION**

As stated in 'The Action Strategy for TB-Free Korea' last March, almost two million Koreans are to be screened for LTBI, using QFT in 2017. Dynex Agility is an automated ELISA analyser developed for QFT; it has...
been used in many laboratories worldwide since its release in 2013. Agility was recently introduced in Korea, and its performance has not yet been evaluated. This study aimed to evaluate the performance of Agility with regards to its laboratory accuracy and speed for QFT.

QFT is a qualitative test and the results are reported as positive or negative according to the IFN-γ level in plasma after stimulation with TB-specific antigens, subtracted from that detected without antigen stimulation.

The assay performance characteristics, based on the manual method, were claimed by the manufacturer to be linear, from 0 to 10 IU/mL. Within-run imprecision (repeatability) CV was estimated to range from 4.1% to 9.1%, with an average within-run CV of 6.6%±0.6% for samples with a concentration range of 0.33 to 7.7 IU/mL. Inter-assay imprecision (between-run precision) CV ranged from 6.6% to 12.3%, with an average CV of 8.7%. The zero IFN-γ (negative control in this study) had a CV of 14.1% in the repeatability analysis and 26.1% in the inter-assay imprecision analysis [11]. In the present study, Agility showed a better precision performance, with a CV ranging from 2.48% to 4.0% and 7.5% to 11.17% in the repeatability and between-run analyses, respectively, at lower (0.25-0.3 IU/mL) and higher (9.5-10 IU/mL) plasma concentrations, compared to the previously described performance with the manual method. The between-day precision over a twenty-day period has not been demonstrated in previous studies; in the present study, relatively higher CV values ranging from 11.93% to 14.84% and 9.69% to 12.83% for low and high concentrations, respectively, were recorded, compared to the repeatability and between-run precision. The CV was shown to be higher for lower concentrations than for higher concentrations in the repeatability, between-run, and between-day precision analyses. As suggested in previous studies, this is plausible because the calculated concentration of IFN-γ is low and variation around a low estimate of concentration will be larger than that for higher concentrations [11]. For the same reason, high CV values ranging from 21.65% to 30.89%, observed in the negative control samples could be due to concentrations being too low and too close to the minimum limit of detection; thus, the quantitative test precision of negative samples need not to be discussed. The linearity regression was sustained from 0 to 10.541, with $R^2$ values consistently above 0.99 for all the tested instruments.

Agility showed high concordance results with conventional methods. The sample showing discordant results had a low TB antigen-nil value near the cut-off; this has been previously reported to be related to high (or frequent) conversion or reversion [5].

While IGRAs have improved specificity compared to TST, concerns have been raised regarding the reproducibility of the tests in settings where repeat testing is necessary [15-17]. Naturally, functional T-cell assays are highly susceptible to variability due to numerous factors at multiple levels, including manufacturing of assay components, pre-analytical processing, analytical testing, and immunomodulation [5]. Therefore, reproducibility is an important consideration that makes it challenging to use a single cut-off value to distinguish between positive and negative results during one-time testing, and to define conversion and reversion in individuals undergoing serial testing.

A systematic review on IGRA reproducibility in 2009, based on a small number of studies, showed that variability was substantial, with magnitudes of within-subject IFN-γ responses varying by up to 80%. Since then, more research has emerged, providing a better understanding of the sources of variability in IGRAs. The ‘total variability’ is the net sum of all types of variability, including manufacturing, pre-analytical, analytical, and immunological sources [5].

Regarding analytical sources of variability, in addition to the precision factors we assayed, the cut-off evaluation could be a concern in laboratory settings for the broader understanding of the near-cut-off results and clarification of the highly un reproducible range around the cut-off value according to the CLSI EP12-A2 guideline. In the cut-off evaluation, the value, 0.35 IU/mL, showed compatibility with the $C_{50}$ value defined in the CLSI guideline. The $C_{50}$±20% range contained the $C_{5}-C_{95}$ interval, suggesting that samples with concentrations ≤0.28 IU/mL and ≥0.42 IU/mL can be expected to yield consistent negative and positive results, respectively.
Samples within C50±20% (range, 0.28 to 0.42 IU/mL), are suspected to show relatively low reproducible results of high inversion between positive and negative. These results are thought to be useful for interpreting results with low TB antigen-nil values, with a possibility of result change in repeated tests.

Agility had a turnaround time of 3.84 hours from sample submission to result reporting per batch. When the first test plate is finished, another plate could be sequentially inserted. The maximum throughput of Agility was estimated to be 464 tests per day, considering the measured throughput per batch and time interval for sequential insertion. However, the actual throughput was measured to be 290 tests per day, which is 2.9 times higher than that of DS2, due to unexpected time lags for completing tests. It inevitably reduced the overall throughput to less than the expected value. However, the throughput was still quite better than that of conventional methods and was thought to be suitable for a large scale test performance.

In conclusion, Agility was shown to possess high precision, concordance with conventional methods, proper linearity in accordance with the manufacturer’s claims, good recovery rates, and rapid performance with high throughput, although this was not maximized as expected. The results of this study would be helpful to many laboratories, considering their need for implementing high-throughput IGRA assays, in compliance with the requirements of the national LTBI screening project.

ACKNOWLEDGMENTS

S. Y. Shin designed and directed this study project. K.H. Seok analyzed data and wrote the manuscript. The authors would like to thank the staffs of the Laboratory Medicine Centre at the Korean Institute of Tuberculosis, Cheongju, Korea. The authors acknowledge that a part of cut-off evaluation was performed as a 2018 research project financially supported by the Korean Association of External Quality Assessment Service.

REFERENCES