Improving turnaround times for HLA-B*27 and HLA-B*57:01 gene testing: a Barts Health NHS Trust quality improvement project

Emma White,1 Nathan Proudlove 2, Delordson Kallon1

ABSTRACT

Among other tests, Barts Health NHS Trust clinical transplantation laboratory conducts two important gene-detection tests: human leucocyte antigen (HLA)-B*27 (‘B27’, associated with the diagnosis of ankylosing spondylitis) and HLA-B*57:01 (‘B57’, associated with prediction of abacavir hypersensitivity disorder). The turnaround time (TaT) from sample receipt to return of results is important to clinicians and their patients but was not monitored. Furthermore, we anticipated an imminent increase in demand from a forthcoming pathology service merger, together with long-term increases with the rise of personalised genetic medicine.

In this quality improvement project, we identified current TaT performance and sources of delay. Over three plan-do-study-act (PDSA) cycles, we tested three change ideas, two involving using IT to remove manual administrative steps and alert us to samples needing processing; both were retained. The other change involved separating out the targeted tests; we judged this not worthwhile with current demand levels, although something to be re-examined when volumes increase. During the project, we reduced mean TaT from 3.8 to 3.3 days and increased the proportion within our 5-day target from 78% to 100%. These have been sustained (at 3.4 days and 97%) for the 3 months following our PDSA cycles and illustrate that reducing variation can be as impactful as reducing the mean.

We conducted this project during the COVID-19 disruption, which reduced demand substantially. We took advantage of this to allow staff to spend time on these improvement activities. Another interesting feature of the work is that during the project, we compared changes in performance on our targeted B27/B57 tests with that on another comparable test as a control, to consider the impact of the general increased attention (the Hawthorne effect).

We found that performance on this control also increased comparably, but then fell away after our project finished, while it did not for B27/B57.

PROBLEM

The Barts Health NHS Trust (BHNT) clinical transplantation laboratory (CTL) (‘the lab’) undertakes organ matching and compatibility testing in support of large renal and stem cell transplant programmes for our Trust and heart/lung and stem cell transplant programmes for Great Ormond Street Hospital. Alongside this work, the lab also provides human leucocyte antigen (HLA) and human homeostatic iron regulator (High FE2+ or HFE) gene testing to a wide range of service users within our Trust and further afield. These tests aid disease diagnosis and avoid patient hypersensitivity reactions in response to certain drugs. The lab provides a service for testing a wide range of HLA genes. For this quality improvement (QI) project, we concentrated on two important tests: detection of HLA-B*27 (‘B27’) and HLA-B*57:01 (‘B57’) genes.

In previous years, substantial improvements have been made in our processes for the renal, heart/lung and stem cell transplant testing pathways. Much of this was possible as these pathways use a laboratory information management system (LIMS), which we manage locally. However, the B27/B57 gene tests use the WinPath LIMS, which is managed on a pathology-wide basis, so these tests did not benefit from the same improvement efforts. The two drivers for the timing of this QI work were: (i) the realisation that for these tests we had not been monitoring the turnaround time (TaT, from receipt at CTL reception to completion (authorisation of the results on the LIMS)). We decided 5 days was a suitable TaT target (see ‘Design’ section), but were unsure whether we were achieving this or not—staff believed that often we were not; and (ii) we anticipated a ~10% rise in B27/B57 test sample numbers in the coming year following the pathology merger between BHNT, Homerton University Hospital NHS Foundation Trust and Lewisham and Greenwich NHS Trust to form the East and South East London NHS Pathology Partnership. As there will be no additional funding with the sample number rise, we will need to absorb this extra demand into existing workflow.
BHNT is the UK’s largest trust, serving a diverse population of around 2.5 million people across East London and beyond. It is well documented that there is an over-representation of black, Asian and minority ethnic patients on both kidney and stem cell transplant waiting lists due to difficulties finding suitably matched donors. Due to our patient demographic, this leads to a busy and high-pressure working environment in the lab, with complex testing for transplant patients managed alongside other diagnostic testing, including B27/B57. Despite these pressures, we are aware of the importance of every test request—from finding a stem cell donor for a patient with leukaemia to confirmation of an ankylosing spondylitis (AS) diagnosis for an anxious patient.

The BHNT vision is ‘to be a high performing group of NHS hospitals, renowned for excellence and innovation and providing safe and compassionate care to our patients in East London and beyond’. This vision is expressed in the WeCare values of being Welcoming, Engaging, Collaborative, Accountable, Respectful and Equitable. At CTL, we are accountable to clinicians who use our service and, ultimately, to their patients. With this in mind, we believed we could improve our B27/B57 test TaT using the BHNT QI approach (WelImprove), which is underpinned by the model for improvement and its plan-do-study-act (PDSA) cycles. This approach uses the model for improvements three systems thinking questions to gain an understanding of the system: “What are we trying to accomplish?” “How will we know the change is an improvement?” and “What changes can we make that will result in the improvements that we seek?” Once ideas have been generated, they are tested via successive PDSA cycles in a disciplined way, learning from the successes and failures in each cycle. The benefits of this method is that it is simple, uses an easy-to-understand systematic approach and has been shown to be highly effective in healthcare settings. It is suited to systems which generate profiles and ease of use. However, ~5% of patients develop hypersensitivity reactions to abacavir, presenting as a combination of fever, rash, gastrointestinal and respiratory symptoms which become more severe with continuation of the treatment. In 2002, this reaction was shown to be associated with the presence of the HLA-B*57:01 allele, which has a prevalence of ~4% in the UK population. As a result, the current standard of care is to ensure that a patient is negative for the B*57:01 gene before starting a course of abacavir.

The key performance indicator (KPI) for our CTL service is TaT, which has long been a marker of laboratory quality; an unsatisfactory TaT can be a major source of complaints about laboratories. Laboratories have a responsibility to patients to produce reliable and precise results within a target TaT so that clinicians can make treatment decisions: lab staff have to balance timeliness and precision. The link between short TaTs and other measures such as length of hospital stay is well documented in the rapid laboratory disciplines such as haematology and biochemistry and has also been the subject of much recent discussion during the COVID-19 pandemic, where rapid TaT of COVID-19 tests has been shown to be crucial in tracking, tracing and containing the spread of the infection. Timely return of results in other areas, such as genetic-based testing, is becoming more important with advances in precision medicine. Personalised therapies based on genetics are a key part of healthcare strategies such as the NHS Long Term Plan.

BACKGROUND

HLAs are proteins on the surface of cells which are crucial for normal immune function. These are a diverse set of proteins with nearly 29,000 HLA alleles identified as of December 2020. Certain HLA alleles have been shown to be associated with specific diseases (such as HLA-B*27 with AS, HLA-C*06 with psoriasis and HLA-B*51 with Behçet’s disease). AS is part of a heterogenous group of conditions with overlapping clinical manifestations known as seronegative spondyloarthopathies. It is a chronic, progressive inflammatory arthropathy which presents as severe back pain and spinal stiffness. There is often a significant delay in the diagnosis of AS, due to the diverse presentation of often generic symptoms such as back pain, along with the lack of a specific laboratory diagnostic test for the condition. The time to diagnosis can be lengthy, as long as 7–10 years, which is associated with poor treatment responses and worse outcomes. As there is no single diagnostic laboratory test for AS, a diagnosis is arrived at through specialist assessment of clinical features. Presence of the B27 gene has long been known to be significantly associated with AS, with a diagnostic sensitivity and specificity of 90%. The current guidelines for diagnosis of AS state that presence of the B27 gene along with two other spondylarthritis features (such as inflammatory back pain, uveitis, psoriasis, enthesitis, etc) is highly suggestive of AS and should be confirmed with an MRI. Therefore, testing for B27 as early as possible can guide clinicians to a more precise evaluation, contributing to earlier diagnosis and so to better patient outcomes.

HLA alleles are also associated with ‘type B’ adverse drug reactions (ADRs): allergic reactions to certain drugs related to host pharmacogenetic factors (rather than drug dose or pharmacology). Examples of ADRs include drug-induced hypersensitivity syndrome in response to abacavir, and Stevens–Johnson syndrome in response to carbamazepine. ADRs cause significant morbidity and mortality and so an understanding of the immune mechanisms which drive the pathogenesis is highly desirable. Abacavir is a nucleoside reverse-transcriptase inhibitor with activity against the HIV. It is part of one of the initial regimens recommended for most patients with HIV, as these regimens have demonstrated durable virological efficacy, good tolerability/toxicity profiles and ease of use. However, ~5% of patients develop a hypersensitivity reaction to abacavir, presenting as a combination of fever, rash, gastrointestinal and respiratory symptoms which become more severe with continuation of the treatment.

REFERENCES

In histocompatibility and immunogenetics (H&I), there has been little published QI work. A PubMed search using search terms “quality improvement” with “HLA” or “Histocompatibility” or “Immunogenetics” only returns one relevant result: a QI project undertaken by the UCLA Immunogenetics Centre on improving electronic reporting of HLA antibody testing results.²⁵ To the best of our knowledge, there is no previous published QI work on HLA genetic testing.

**MEASUREMENT**

Figure 1 shows the workflow and our metrics. Our KPIs were the twin metrics of mean TaT (KP1) and % within our TaT target of 5 days (KP2). As noted, we expect demand to increase in future, but COVID-19 disruption caused a major reduction as many outpatient and GP appointments were cancelled or moved online during the pandemic. We monitored demand as a balancing metric (BM). During the project, we also introduced two (internal) process metrics (PM) for monitoring change impact (PM1 to monitor the percentage of tests that were included on the next available Luminex run, and PM2 to monitor test to authorisation time), plus an individual test sample progress monitoring measure (time in process so far (TiP), so we could identify samples at risk of breaching the target).

We collected two sets of TaT data as baseline measurements: pre-COVID-19 (September–November 2019) and during COVID-19 (June 2020), to assess baseline TaTs during both previous workloads and the reduced (‘new normal’) volumes.

The online supplemental figure shows the five metrics over all periods of interest. The KPI and PM graphs are in statistical process control (SPC) format, showing the mean underlying levels, the three sigma process variation limits for ‘common cause’ (random) variation and potential non-random datapoints (‘special cause’ variation).³ In addition, we show a ‘results’ section in the supplemental file.

**DESIGN**

We decided on a 5-day TaT target (KP1) for a number of reasons. First, this is the target which other NHS H&I laboratories in England are using for these tests and therefore brings us in line with the national benchmark. Second, this is relatively fast compared with TaTs for our other tests (typically 14 or 21 days) and would reduce potential further delays to patients starting treatment.

---

**Figure 1** Process map. BM, balancing metric; KP, key performance; PM, process metric; TiP, time in process so far.

---

especially for abacavir treatment for HIV (which depends on the B57 test).

We decided on an additional target (KP2): at least 90% achievement of the 5-day TaT, to be consistent with performance measurement reporting for the other tests we provide.

The lead author (EW) led this QI project, starting with a socially distanced CTL team meeting in June 2020 to discuss the project; staff were enthusiastic and suggested where delays were occurring in the B27/B57 pathway along with ideas for improvement. Based on this meeting, we drew a process map of the pathway (figure 1), highlighting the areas causing delays. We then formed a subteam of scientists and lab technicians to test the changes. This smaller team met after each PDSA cycle (every 3 or 4 weeks) to discuss the results from the current cycle and plans for the next.

The main constraint anticipated at the start was a block to major changes to the LIMS, since the pathology ICT team were working on a pathology-wide project and so had frozen any major changes for individual departments. Therefore, our change ideas were limited to small changes to the LIMS or other process changes we could make within our own team.

**STRATEGY**

In our project time-period, we undertook three PDSA cycles (table 1).
Plan-do-study-act 1

In Pareto analysis, categories of causes of an outcome are displayed in order of frequency of occurrence to help focus QI. This analysis showed that the most frequent category associated with breaches of the TaT target was delay between sample receipt and sample set up on a Luminex run (figure 1). This, then, was the focus for our first intervention. Root cause analysis showed that a particular problem here was failing to include a sample on the next run. We therefore established PM1 to monitor the percentage successfully included on the next available run, and so pick up the impact of intervention. The baseline level was only 70% (figure 4).

Our change was to replace a manual, paper-based worklist-generation process with a faster and more reliable automated LIMS query which generated an Excel worklist. The results were dramatic improvements in PM1 to 97% on the next Luminex run (figure 4) and in overall TaT, with 92% of samples now within 5 days TaT (figure 3). In addition, feedback from the staff was very positive. This was adopted as a permanent change, and we carried on monitoring PM1 for 3 more months (reliably achieving 100% of samples on the next run).

This change, though, had the unintended consequence of requiring us to overhaul our sample set up methods. Previously, the paper worklist had also been used to record important assay information (dates of sample set up, analysis and authorisation and reagent lot numbers) and was also used to track where each sample was within the system. When we eliminated the paper setup sheet in this PDSA, we realised this extra information had to be recorded elsewhere. To solve this, we created electronic worksheets to record the run setup/reagent information and an electronic master sample log where we could track each sample in the system. Feedback from staff about these follow-on changes was also very positive.

Plan-do-study-act 2

Despite the dramatic improvement from PDSA1, we wanted to continue to look at other areas of our process, so that when sample numbers return to their pre-COVID-19 levels we would have a robust system in place to maintain TaT performance. The next identified delay highlighted on the process map was between sample test and authorisation steps (figure 1). We set up PM2 to monitor the time between these stages and established that the baseline performance on it in the period prior to PDSA2 was 2.0 days.

To reduce this delay, in the analysis and checking stages we separated B27/B57 tests from the renal and stem cell transplant tests (since these latter ones take longer). This reduced PM2 to 1.8 days (mean time between testing and authorisation), but the variation increased. Although this resulted in a further increase in the percentage of samples authorised within 5 days (to 98%, figure 3), it also resulted in an increase in the number of sets of samples at different stages of analysis and checking. Feedback from staff was that managing this required closer communication between each other. We decided the sample volume at the time did not warrant this extra complication, so the decision at the ‘act’ phase was to not retain this process change.

Plan-do-study-act 3

This built on the success of the automated system in PDSA1 and the electronic data recording for B27/B57 samples created as a follow-on. We decided to try to use this new tracking system to automatically calculate each sample’s time-in-process so far, indicated as TiP in figure 1. We then set the IT to generate a daily automated ‘amber alert’ email warning about samples booked in 2+ days previously (so currently 3 days or less from 5-day TaT target limit—and so at risk of breaching it). This email was received by the clinical scientists every morning to enable the scientist in charge of the day-to-day workload to deploy staff as required to prioritise analysis/checking and avoid breaches. This change
Table 1  Improvement cycles

<table>
<thead>
<tr>
<th>PDSA cycle</th>
<th>Plan/Prediction</th>
<th>Do</th>
<th>Study</th>
<th>Act</th>
<th>Time required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline ('new normal')</td>
<td>Remove need for a clinical scientist to create paper B27/B57 worklist (manual transcription step) for Luminex setup.</td>
<td>Create automated query to produce list of samples received the previous day, saved as Excel spreadsheet on shared server. Sample list then make available first thing in morning to the lab technician setting up Luminex run.</td>
<td>KP1: mean TaT=3.8 days. KP2: % of TaT≤5 days=78%. PM1: % on next run=70%. PM2: mean test→authorised=2.0 days.</td>
<td>Worthy improvement. Adopt as permanent change. (S shown to be reliably sustained over the subsequent 3 months: all samples added to next available Luminex run.) Create new electronic master sample log.</td>
<td></td>
</tr>
<tr>
<td>PDSA1</td>
<td>Separate analysis and checking of B27/B57 (quicker) from renal/stem cell samples (slower).</td>
<td>Introduce new parallel process.</td>
<td>PM1: 97%—success. Only one instance of a sample not being set up on the next available run (figure 3). KP1: 3.5 days. KP2: 92%. Unexpected: lose recording of assay information.</td>
<td>3 weeks</td>
<td></td>
</tr>
<tr>
<td>PDSA2</td>
<td>Improve tracking of B27/B57 samples so analysis/checking/authorisation not missed in daily work planning.</td>
<td>Add B27/B57 tracking to existing electronic dashboard and ‘amber alerts’ system based on time in process so far (TIP). Create automated query to generate ‘amber alert’ list of all B27/B57 samples booked in &gt;2 days previously (TIP &gt;2 days). Email list to all clinical scientists first thing every morning.</td>
<td>KP1: 3.3 days (not material change), but variation lower (good). KP2=100% all samples had TaT 2, 3 or 4 days; no samples breached TaT target. TIP is a useful tool for workflow progress planning.</td>
<td>4 weeks</td>
<td></td>
</tr>
<tr>
<td>PDSA3</td>
<td></td>
<td>Successful: more capable process. Adopt as permanent change, retain TIP.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

KP, key performance; PM, process metric; TaT, turnaround time.

was very successful: no samples breached TaT in the 4-week PDSA period and staff feedback was positive. Therefore, the change was retained.

RESULTS
The results are illustrated in figures 2–4 (and the online supplemental figure). Our main outcome measures were mean sample TaT (KP1) and the percentage authorised (completed) within 5 days (KP2). TaT had reduced from the pre-COVID-19 levels (mean TaT 5.1 days) to a new baseline (mean TaT 3.8 days) as the sample volume (BM) decreased during the COVID-19 pandemic. Through this QI project, the TaT dropped further, reaching a mean TaT of 3.2 days with PDSA2. The percentage completed within 5 days (KP2) increased rapidly as a result of the interventions, and we achieved our aim of at least 90% of samples authorised within 5 days by the end of PDSA1 (figure 3). The graphs also show that the performance
has been sustained for the 3 months following our PDSA cycles, while the demand has increased (though not yet to pre-COVID heights). There have been a few potential special cause outliers, which we investigated and found to be due to various rare events (such as rare allele samples or unusual results requiring repeats) and one instance of staff missing the electronic worklist, from which we have learnt. The mean TaT (KP1, figure 2) for the post-PDAs period is 3.4 days, and (the above events aside) it is ‘capable’ of reliably meeting the 5-day target since the upper of the red three sigma process variation limits does not exceed the target. Similarly, KP2 (figure 3) is 97% of individual TaTs within 5 days, and just-about capable of reliably meeting the 90% target. The special cause here was triggered by the incident noted above where a work-list was missed.

The change which had the largest effect was PDSA1, the replacement of the manual paper worklist with an automated sample list, available first thing each morning so that the lab technicians did not have to wait for it to be created by the duty scientist. This was very successful in ensuring that every sample went on the next available run (figure 4).

The SPC graphs in figures 2–4 illustrate that targeting a source of delay can reduce variation significantly, and that this can have a dramatic impact on the proportion within a target (so the success rate) even if the impact on the mean time itself is relatively modest; variation can matter as much or more than the mean. The graphs also show that the changes made have had a sustained effect on the TaT (figures 2 and 3).

LESSONS AND LIMITATIONS

The most significant limitation in this QI project is that the changes were implemented at a time when sample numbers were materially lower than usual because of the COVID-19 pandemic. We mitigated this limitation by obtaining two baseline data sets: both pre-COVID-19 and during COVID-19. Although this meant we were not testing the changes at normal sample volumes, the reduction in volume was beneficial as it allowed us time as a lab to examine our systems closely and work on change ideas as a team. On the other hand, our team, as across the NHS, had to adapt to incorporating some remote working during 2020. This meant that the entire team was unable to meet in person during this QI project. The PDSA2 experiment (separating workflows during analysis and checking) revealed that that change required additional staff communication and coordination. The reduced routine staff contact during COVID-19 may have contributed to our staff judging that this requirement made the change not worthwhile at that time.

Another potential limitation of the changes made is the confounding hypothesis that the improvement in TaT could be due to increased staff attention during the QI: the Hawthorne effect. Since B27/B57 samples are a small part of our workload, and one on which we never reported our KPIs, these samples were always lowest of staff priorities. They instinctively focused on our transplant work—for which we are required to have results ready for multidisciplinary team meetings and for which we regularly examine our KPIs.

To examine this potentially confounding effect, we compared the within-5-day performance of B27/B57 with that of our HFE gene test (a similar standalone test for which, historically, we also had not examined TaTs). The baseline data for the HFE TaT were similar to B27/B57: 55% authorised within 5 days during September–November 2019 and 79% in June 2020. During this B27/B57 QI project, the HFE performance also increased to 96% and 92% in July and August 2020, respectively, despite no changes having been made to that pathway. However, following the end of the increased attention stimulated by the QI project in the lab, the performance for B27/B57 has remained high (98%, 97% and 95% for September, October and November 2020, respectively; figure 3), whereas for HFE it has been decreasing (97%, 87% and 83%). This suggests the Hawthorne effect may have impacted performance during the PDSA cycles, but it is not responsible for (all) our new performance level for B27/B57. It is, though, a reminder for future QI projects of this potential confounding effect, the importance of demonstration sustained change and the value of ‘control’ comparisons where possible.

A potential limitation of the change implemented in PDSA3 is that staff focus on expediting samples about to breach could adversely affect progress of other samples. Therefore, an adaptation to consider could be ‘time buffer management’ as has been used in accident and emergency and inpatient wards. 27, 28 This would use the TiP measure to predict the time of completion and compare this with the required time of completion, protected by a time buffer. For each sample, the status of its time buffer would determine its progress status, which would determine its priority, colour-coded in the dashboard system. A sample slipping towards the red zone (predicted to be in danger of using up all its time buffer and so at risk of breaching the target TaT) could then be targeted, particularly at the two main stages where delays were identified by our Pareto analysis: sample set-up and result authorisation. This information could also be used for further analysis, for example, regular investigations of root causes for samples entering ‘red’ status.

CONCLUSION

NHS laboratories are increasingly being asked to provide more tests and services for patients, with drivers including the demand for more genetic testing as the NHS moves towards personalised medicine. However, the corresponding additional funding and staff is often not available. A parallel pressure is the need for earlier availability of results to enable clinicians to make earlier treatment decisions and so increase benefit to patients.
Consequently, increasing automation and reducing staff hands-on time in the labs is important.

We have demonstrated that it is possible to examine processes in a genetic testing pathway, make straightforward process improvements and materially improve TaT performance through a QI process. We were pleased to fairly quickly achieve our aim of achieving at least 90% of samples being reported within 5 days of receipt, establish that this change was sustainable and start monthly monitoring of TaTs. We hope this project will demonstrate the potential to other H&I laboratories; further work is underway to disseminate this QI project both within the Trust and to H&I laboratories nationwide.

Acknowledgements The authors would like to thank everyone at the Clinical Transplantation Laboratory at Barts Health NHS Trust for their contribution and work on this project. EW conducted this work as part of the QI unit of the Higher Specialist Scientist Training Programme (HSSST), run by Alliance Manchester Business School of the University of Manchester and the Manchester Academy for Healthcare Scientist Education (MAHSE), on behalf of the National School of Healthcare Science, part of NHS Health Education England.

Contributors EW designed the work in consultation with DK. EW performed the collection, analysis and interpretation of data. EW and NP drafted the work. All authors reviewed the draft. DK is the guarantor of the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been vetted by any third party. The content is included in the article for informational purposes only and BMJ can not accept any liability for any errors or omissions.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD
Nathan Proudlove http://orcid.org/0000-0002-1176-8088

REFERENCES