Innovating for Improvement

[Listen, Learn & Improve: Using language analysis to interpret and act on written patient experience feedback for near real-time patient benefit]

[Imperial College Healthcare NHS Trust and Imperial College London]

The Health Foundation

About the project

Project title:

Listen, Learn & Improve: Using language analysis to interpret and act on written patient experience feedback for near real-time patient benefit

Lead organisation:

Imperial College Healthcare NHS Trust & Imperial College London

Project lead(s):

Erik Mayer

Stephanie Harrison White

Bob Klaber

Joshua Symons

Mustafa Khanbhai

Dave Manton

Kelsey Flott

Jamie Spofforth

Contents

Part 1: Abstract

Understanding patients' experience of health care is central to the process of providing patient-centred care and is a fundamental pillar of health care quality. Consistent use of patients' feedback on their experiences has however been inhibited by the unstructured format of the data (it being free text narratives) and therefore the manual review needed to extract insights. In this project we used semi-automated analysis, using Natural Language Processing (NLP), to analyse these free text narratives using quantitative methods to extract patients' experiences of their care, in near-real-time, and subsequently used this feedback in a quality improvement (QI) effort.

Using a co-design approach, we ensured that challenges were addressed and risks mitigated with regular meetings and updates. We demonstrated that to establish continuous change, QI aligns with the interests of front-line staff and key stakeholders. Some of the barriers encountered with multi-disciplinary team working such as competing demands forced rationalisation of efforts and in some cases consciously relinquish other priorities.

Our approach of rapidly, and continuously, analysing patient feedback using NLP meant that changes resulting from this information occurred in a timelier manner and then implemented into the regular workflow of the staff on the ward. This reinforced staff efforts in addressing and sustaining change for patient benefit and ensured that gaps in patient experience were identified as early as possible and acted upon quickly. Ultimately, this project revitalised the benefits of hearing from patients in their own words by holding narrative data to the same standard of scientific rigour that is already applied to conventional patient experience surveys.

Part 2: Progress and outcomes

There is a substantial evidence base in the UK to suggest that patient experience feedback remains under-used for quality improvement (QI) in healthcare. Free-text information within the surveys holds richer information and staff can relate to this feedback as the comments seem more real than Likert scale responses. However, the ability to interpret and analyse the free-text portion falls short in most healthcare services. This is the result of many factors like the timeliness of feedback, its ability to link to other quality indicators and the resource intensity required to manually extract the necessary information. Asking for and collecting patient feedback without then acting on it presents ethical implications, as patients put time and energy into providing this meaningful information, but this then fails to drive change. The information contained within the free-text patient feedback complements structured feedback and helps contextualise responses to closed questions like those in the National Patient Survey Programme. Automated or semi-automated analysis of free-text patient feedback using Natural Language Processing (NLP) allows human language input (free text narratives) to be analysed using quantitative methods, thereby streamlining patient experience improvement efforts in near-real time.

In this Innovating for Improvement project we used the capabilities of NLP and machine learning (ML) to successfully, reliably and continuously extract coded data from free-text patient experience feedback and use this information to develop visualisations for frontline staff to then act upon it as a QI initiative.

Using the government mandated Friends and Family Test Survey, we collated retrospective free-text patient feedback data, from January to July 2017 across four healthcare services (Inpatients, Outpatients, Accident and Emergency [A&E] and Maternity) at Imperial College Healthcare NHS Trust (ICHNT). The free-text comments within individual responses from this initial dataset were used to build and test the ML algorithm. KNIME analytics platform, was used to explore the ability of supervised learning to produce meaningful insights into patient experience data. KNIME is an extensible program utilising workflow-based structures and open-source plugins to enable easy access to data extraction, transformation and advanced analysis, including NLP and ML. KNIME was chosen because of its free and open-source nature, therefore maintaining our aspiration to scale our project at other NHS Trusts. This software resided within a Trust server and the patient experience survey responses were downloaded onto the server at regular intervals. This workflow meant that patient comments could be analysed in near real-time as they were submitted. Free text comments were examined in response to the questions: "What did we do well?" and "What could we do better?". Ten percent of the comments were manually coded into themes according to the NHS Patient Experience Framework (PExF) and the sentiment of each comment was rated as positive, negative, or neutral. The patient experience lead and a lay representative coded five percent of the same comments and inter-rater agreement was assessed.

Supervised learning, where input is mapped to output labels, was used by training six ML algorithms to classify comments and assess their performance. We did also attempt to use an unsupervised ML learning approach, where the inherent structure of the data is learnt without providing explicit labels. This in theory would identify topics based on similar words used within the free text narratives, rather than assigning them to the pre-defined themes of the patient experience framework. This second approach, however, was not technically feasible as the average character count of the free-text narratives was too small to detect any differences in words to generate different topics.

We completed the supervised learning approach and configured each output instead of using the default settings at individual pre-processing steps to ensure the learning data set

would not compromise the accuracy of the ML model. 131,946 free-text comments were analysed, from four hospitals at ICHNT. Inter-rater agreement was 0.852 and 0.952 for PExF theme and sentiment respectively in Inpatients p<0.001, with similar results in the other services. Of the six ML algorithms tested, Support Vector Machine (SVM) learner produced the highest accuracy to the question "What did we do well" in Inpatient, Outpatient, and A&E settings; 70.5%, 66.9% and 75.7% respectively, with similar accuracy to the question "What could we do better". The ML algorithm processed the comments much faster into themes than manual review. The time spent on manual processing was four days compared to 15 minutes using the ML algorithm, p<0.001.
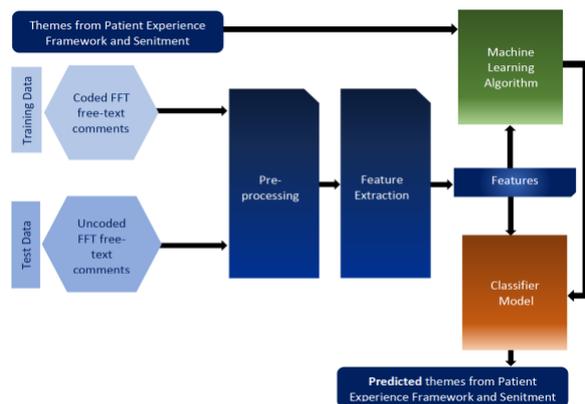
The flow diagram demonstrating the steps we used to develop a machine-learning model to predict sentiment and themes.

To translate the findings from the free-text analysis into actionable interventions, we conducted a series of focus groups to develop a dashboard to enable staff to visualise the NLP output, followed by coaching to begin the QI work using the Institute for Healthcare Improvement's Model for Improvement. Engaging staff in problem solving during a series of focus groups was vital to this project, as their role in delivery health care is central to patient experience, so they needed to determine the design and look of the visualisations and how it can be subsequently improved. Therefore, using an iterative approach, a dashboard was created using a data visualisation tool, Tableau, where the output from the NLP free-text analysis was taken from KNIME and put into Tableau.

The dashboard was presented with free-text comments split by themes and sentiment in near real-time. The consensus following the focus groups favoured formatting our experience visualisations using the same format as other visualisations used within the Trust, for example, our Patient Safety dashboard consists of the top five safety incidents using a bubble chart. This was particularly important as our staff raised the issue that the pre-existing format of patient experience survey reporting used too much technical language, which required training in data analysis and statistics to facilitate its full understanding to then use the results appropriately. Therefore, adopting a new visualisation that was different to the format of the Patient Safety dashboard may not result in any changes, as staff would find the data too far removed from their day-to-day services, and because it would not present the themes felt to be important, as a result find it challenging to translate the NLP output into actionable interventions. We used Tableau as our visualization tool, and the software was downloaded within the Trust server and all staff involved in this proof-of-concept study were given access to the server. We recruited frontline staff from A&E and Inpatient services who were coached by the QI team to become champions and demonstrate that the visualisation can instigate QI work. We started using

a prospective dataset for the QI project, so that changes in comments over time can be assessed in response to this project and therefore any subsequent QI work would be done in a timelier manner.
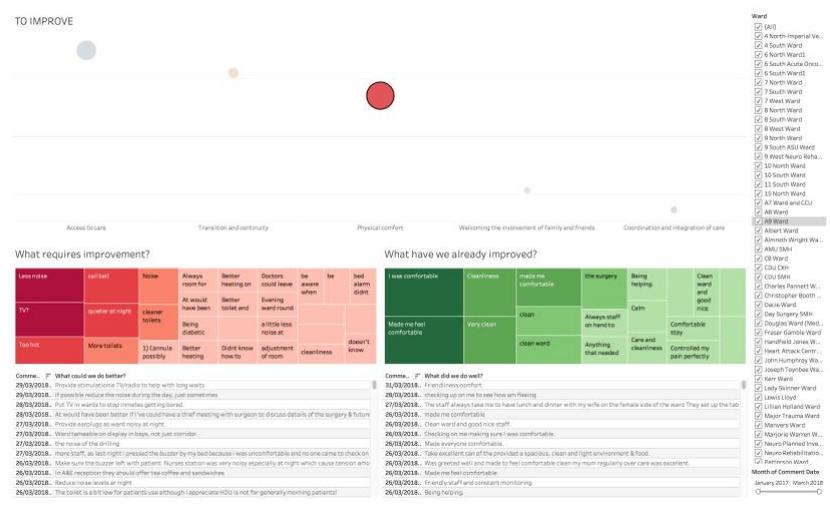
A working visualization demonstrating comments separated by themes according to negative sentiment in descending order. The size of the circle relates to the number of comments within that particular theme. In this example, all the negative and positive comments associated with 'physical comfort' are for all Inpatients for the month of January 2017 to March 2018, inclusive.

We used a bespoke QI framework, 'Project on A Page' that was developed by the QI team at ICHNT. Five staff from each healthcare setting (A&E and inpatients) were involved in a 30-minute brainstorming session whereby they looked at the NLP output visualisation for the months of July to October inclusive. The top three themes based on the most negative sentiment were chosen and the individual comments below were examined in more detail to understand the granularity of the comments in each theme. A SMART objective (Specific, Measurable, Achievable, Realistic and Timely) was created. For this particular project, the objective was, 'To improve patient experience using the NLP dashboard to generate small tests of change based on negative patient comments and suggestions for improvement, and complete three PDSA cycles'. Outcome, process and balancing measures were developed to enable staff to know that the change is an improvement, and ideas for changes were discussed. Of the engaged staff members, all were delegated to be involved in certain aspects of the tasks to ensure the action was followed through and there was a point of call if required. Many of the change ideas were generated by patients directly and this allowed us to test ideas suggested by patients. Also, of interest in improvement work are the less frequently reported comments and themes, as this information is easily overlooked when looking at other sources of high volume patient feedback. This will be part of further QI work.

Implementing change using 'Project On A Page', enables staff to systematically understand and evaluate the comments raised by patients, based on the NLP visualisation. It demonstrates that meaningful change can be generated in this way and encourages 'buy in' from staff as the patient comments seem more real to them.

This

Part 3: Cost impact

This section is intended for you to review the cost impact of your project – giving an indication of whether the intervention is cost saving, cost neutral or requires ongoing investment.

Listen, Learn and Improve is a project based across multiple services and geared towards improving the use of patient feedback across providers. The improvements generated from this work were contingent on an initial investment to develop the ML algorithms and solidify the method for quality improvement.

For this reason, the Innovating for Improvement award has funded the initial investment of a larger vision to improve the use of patient feedback across the NHS. The remainder of the work can now be conducted elsewhere, provided there is a server on which to store data. It can be run with a great degree of local ownership while maintaining academic validation.

In terms of an economic evaluation, this was not conducted during this phase of the work. In the future, however, it would be prudent to undertake an economic evaluation of how our newly developed process for using patient feedback in more direct attempts to improve care compares with existing methods. There is evidence to suggest that feedback collection exercises in the NHS are costly endeavours. This investment is significant and it is important that it be worthwhile and a contributing factor to meaningful QI. This work helps ensure that data from costly feedback collections is used. In many cases this work can also reduce the staff time required to manually sift through patient feedback without any support to identify patterns.

Costs for this project were calculated based on the staff time needed to complete the work, as well as the equipment required. In the case of ICHNT, this included purchasing a server. A dedicated server was used for this project due to the development nature of the NLP algorithm and visualisations. Now that the workflow has been established, processing demands could easily be incorporated into a Trust's IT infrastructure, depending on existing local computing infrastructure, thereby avoiding this server cost. Other costs incurred in this project included compensation (expenses) for staff and our patient co-investigator. When applying the ML techniques to free text patient feedback in other Trusts, it is likely that costs would mainly relate to staff time to understand the process, coordinate the feedback of results to staff and maintain responsibility for monitoring whether or not trends in feedback are changing. The key limitations or concerns for spreading or scaling this work would revolve around the availably of staff to take forward the work and provide local championship and the capacity of servers for data storage.

As intended at outset of this project, we believe our work can be delivered at relatively minimal expense at other healthcare providers and may even provide a cost saving depending on current local arrangements for collecting and processing only 'structured patient feedback' using external providers.

Part 4: Learning from your project

We established a steering group that allowed for co-design, which consisted of key stakeholders from four departments; Patient Experience and Quality Improvement based within Imperial College Healthcare NHS Trust, Patient Safety and Translational Research Centre, Big Data and Analytical Unit (BDAU) and Centre for Health Policy. Members from each department met to discuss and coordinate various aspects of the project. The diversity of expertise within the team provided insight into machine learning and NLP as well as NHS service delivery and the complex nature of local NHS structures. Furthermore, we conducted a process mapping exercise that unravelled how patient feedback data flowed within our healthcare organisation. This helped us understand pain points and identify and prioritise areas for improvement. In its current format, we noted that the slow (and uneven) progress from ideas to implementation of patient feedback, and the way QI projects have been adapted with an attempt to fuse with other quality metrics did not stimulate the actual changes, challenged our own morale and confidence about the impact of the work. Pain points included frontline staff not having a say in how patient experience data is presented, software difficult to navigate resulting in poor user experience, lack of time and resources to look through patient experience comments in a timely manner and insufficient knowledge in driving QI projects based on patient feedback.

Working as a collaborative group enabled shared decision making, with patient and public involvement and engagement (PPIE) at key stages throughout the project moulded the project to be patient-centric. In addition to a lay representative who was part fo the sterrting group, we presented our proposal to the Research Partners Group at the Imperial Patient Safety Translation Research Centre. The RPG positively impacted our research project, we learned about PPI and the value of it and RPG members also benefited from their participation. Using this approach we noted that there was equality of legitimacy and value in inputs from all those involved, whether suggestions entail large- or small-scale changes. During the initial stages where our protocol was being refined, feedback from all individuals from the PPIE group was invaluable. For example, a member of the group who had experience with customer services in telecommunication advised us to develop our QI strategy based on other customer experience driven industries, and another member from the Department of Education advised us to ensure a lay member helps with manual coding.

There was an initial delay in acquiring the server within the Trust, which was needed to ensure comments were collected and resided within the Trust data warehouse. The software KNIME and Tableau were also downloaded on the server. The benefit of purchasing a server ensured that all data was analysed and reported within the Trust IT infrastructure, and only individuals with Trust contract were able to access it. However the delay was mitigated as we commenced manual coding whilst the server was being set-up by the IT team. The infrastructure and clinical engagement were already in place within the Trust and mechanisms were developed to allow regular data, information and knowledge sharing. The BDAU worked closely with QI and service leads to ensure findings were meaningful to them. Referring to the timelines as per the Gantt chart ensured that timeline was adhered to and in circumstances where delays were anticipated, tasks were prioritised and the wider team made aware.

Due to the government mandated Friends and Family Test all healthcare organisations within England have been required to collect data on patient experiences. Therefore this retrospective data was readily available within our Trust, and after procurement of the server we were able to start work on the project, i.e., manual coding with no significant delays. During the latter part of this project, liaising with the QI coaches to deliver weekly QI coaching to staff and ward teams to understand and lead QI helped overcome some of the cultural inertia to embrace change. The use of collaboration to address challenging themes that many healthcare settings have in common, such as physical comfort or access to care enabled organisational learning.

After an initial learning curve with respect to developing the NLP and ML algorithms, the next hurdle was ensuring staff would engage with the visualisations on a regular basis. Preliminary focus groups revealed that staff members found free text comments more helpful and revealing than quantitative responses, as they provided the context and detail required for them to learn from and act on patients' views. During the focus groups almost all staff were enthusiastic and keen to use the new dashboard, as this was driven by feedback from frontline staff. The consensus was that the staff would use and integrated this dashboard into their daily routine. Changing the mindset of staff with regards to the importance that patient experience should be examined together with clinical effectiveness and safety was vital in ensuring that patient experience is not considered as remote adjuncts to clinical work. Therefore, whilst adopting a new visualisation, feedback from staff ensured that the final dashboard would be bespoke; this was arguably the most important factor for transformational change to occur. Despite the initial positive feedback, we encountered barriers from staff due to lack of time or resources. We noticed that the staff who were involved in designing the visualisation tool were more engaged than those who were not aware of this project but were approached to help in the implementation phase. We refer to clinical engagement, as to whether staff were actively contributing 'within their normal working roles to maintain and enhance the performance of the organisation which itself recognises the commitment in supporting and encouraging high quality care'. Co-design with frontline staff and other key stakeholders was crucial for spread of this innovation. If other Trusts just implemented our innovation without any co-design around the QI aspect, there would be a real risk that this innovation would never properly be used.

Part 5: Sustainability and spread

We utilised an open source platform that is now embedded within the Trust server with regular data feeds from the patient experience team. Using an iterative process we have streamlined the flow of the experience reports in near real-time, resulting in timely accessibility for staff to drive QI. This has been rolled out into all four healthcare settings; Inpatient, Outpatient, Accident and Emergency and Maternity. The four lead champions have begun to disseminate their knowledge to other frontline staff in their respective departments with the help of the QI team. This ensures that the dashboard continues to be used. With their consent, we have begun to audit the frequency of log-in and the duration an individual staff member spends on the dashboard. The outcomes of the QI projects using the Project on a Page, enables the patient experience team and QI team to understand and monitor changes in patient experience over a given timeline. Through the Communications team at the Trust, and advertising campaigns on the Trust's Intranet, we have attracted interest from numerous staff members and other departments, for example, radiology and physiotherapy who are keen to challenge their existing patient experience reporting strategies and utilise our innovation for direct near real-time patient benefit.

Our collaboration with the BDAU applies advanced analytics to healthcare data for the purpose of direct care returns, while adhering to best-practice in information governance, data linkage, data collection and recognises the needs of health care professionals and patients for analysis and interpretation. The unique Data Processing agreement allows pseudonymised data to be held to an equivalent standard as fully identifiable data (ISO 27001:2013/NHS IG Toolkit Level 3 (EE133887), and is integral to facilitating meaningful analytics to support clinical QI, reduce clinical variation and improve population health.

To date, we have had the opportunity to present our work to the NHS England Insight and Feedback team, Nursing and Midwifery Conference at Hammersmith Hospital, Share and Spread Improvement Conference at ICHNT, Executive Directors of Nursing Conference in Cambridge, Patient-Reported Outcomes, Value and Experience Centre (PROVE) at Harvard Medical School and NIHR Patient Safety and Translation Research Symposium held in London. Our work has been accepted at the International Digital Health and Care Congress 2019. We have also received interest for a transatlantic collaboration with the Mental Health Innovation Lab New York City Department of Health & Mental Hygiene. We are working closely with CareOpinion by applying our algorithms to the free-text data and comparing the output with the manually populated tags. This will be followed by developing new themes that do not encompass the Patient Experience Framework by using a different ML approach, unsupervised learning.

Through NIHR Clinical Research Network Adoption (Portfolio-IRAS 234218), this solution is being deployed across other Trusts. In collaboration with NHS England Insight & Feedback Team, the application of the NLP tool to primary care patient feedback is being undertaken. Analysis of NHS choices data, with its national coverage, is providing 'Hospitality industry' style visibility of patient experience and understanding of system behaviours. A visiting fellow programme has been initiated with the Patient-Reported Outcomes, Value and Experience Centre at the Brigham and Women's Hospital, affiliate of Harvard Medical School in Boston who are applying the NLP to US healthcare data to gain a deeper understanding from an international perspective on the translational impact of this

work. However, new healthcare services who are embarking on implementation of similar innovation must exercise caution particularly encouraging co-design during dashboard creation and staff engagement.

Appendix 1: Resources and appendices

vom Eigen, K.A., et al., *Carepartner experiences with hospital care.* Med Care, 1999. **37**(1): p. 33-8.

Hewitt-Taylor, J. and C.S. Bond, *What e-patients want from the doctor-patient relationship: content analysis of posts on discussion boards.* J Med Internet Res, 2012. **14**(6): p. e155.

Fullwood, C., et al., *Patients' experiences of shared decision making in primary care practices in the United kingdom.* Med Decis Making, 2013. **33**(1): p. 26-36.

Rathert, C., N. Huddleston, and Y. Pak, *Acute care patients discuss the patient role in patient safety.* Health Care Manage Rev, 2011. **36**(2): p. 134-44.

Coulter, A., et al., *Collecting data on patient experience is not enough: they must be used to improve care.* BMJ, 2014. **348**: p. g2225.

England, N., *The Friends and Family Test*. 2014, NHS England.

Robert G, C.J., *What matters to patients? Developing the evidence base for measuring and improving patient experience.* 2011: London: The King's Fund.

Gleeson, H., et al., *Systematic review of approaches to using patient experience data for quality improvement in healthcare settings.* BMJ Open, 2016. **6**(8): p. e011907.

Friedman, C., et al., *Automated encoding of clinical documents based on natural language processing.* J Am Med Inform Assoc, 2004. **11**(5): p. 392-402.

Lilley, E.J., et al., *Measuring Processes of Care in Palliative Surgery: A Novel Approach Using Natural Language Processing.* Ann Surg, 2018. **267**(5): p. 823-825.

Yim, W.W., et al., *Natural Language Processing in Oncology: A Review.* JAMA Oncol, 2016. **2**(6): p. 797-804.

Please attach any leaflets, posters, presentations, media coverage, blogs etc. you feel would be beneficial to share with others.

A visiting fellow programme has been initiated with the Patient-Reported Outcomes, Value and Experience Centre at the Brigham and Women's Hospital, affiliate of Harvard Medical School in Boston who are applying the NLP to US healthcare data to gain a deeper understanding from an international perspective on the translational impact of this work.
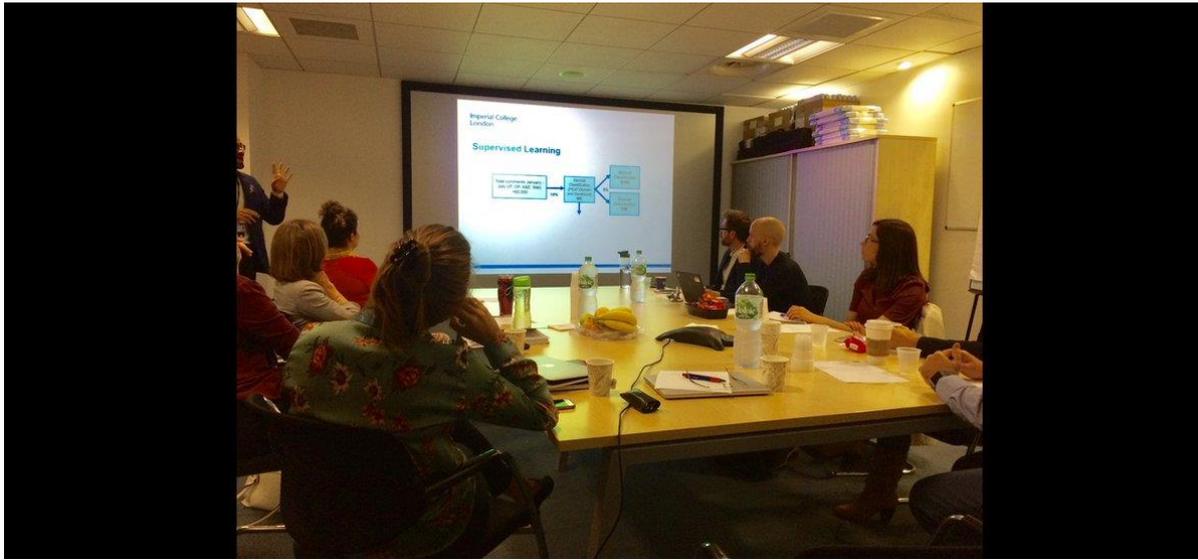
Presenting to NHS England Insight and Feedback Team, October 2018.

Poster presentation at the annual Patient Safety Translation Research Centre Symposium, November 2018.