

Ensuring the Safety of Babylon Products and Services

Introduction from our clinical and scientific team

“Babylon believes in putting an accessible and affordable health service in the hands of every person on Earth. As senior clinicians and scientists at Babylon, we take our responsibilities to provide safe and effective care extremely seriously. Creating services that are available anytime, anywhere at the touch of a smartphone button is core to what we do, but even more important is the ongoing testing, quality improvement and external assurance that we undertake to ensure the clinical safety of all elements of our work.

Our job is not just to build and provide great services across the world, it is to set the climate and culture across the 600 people who work for Babylon, including 200 GPs, embedding patient safety as a priority in everything we do. We believe in transparency, we welcome the feedback from the global community of clinical colleagues and the general public, and above all we know that no matter how a service performs today, it can always be better tomorrow. By setting out our approach, we hope to bring even greater transparency to our work, to help raise the bar on patient safety, and to dispel a few myths along the way.”

Dr Mobasher Butt, Babylon Medical Director

Mobasher is a practicing GP, with expertise in Patient Safety and Quality Improvement. He led mobile health research on a global scale for the World Health Organisation, was Clinical Advisor to the CMO for England and Quality Improvement lead for the BMJ.

Dr Keith Grimes, Babylon Clinical Innovation Director

Keith is an NHS GP with 18 years' primary care experience across General Practice, Urgent and Out-of-Hours care, and Walk-in centres. His special interest is Digital Health, Innovation, and Clinical Safety, and he is an NHS Digital trained Clinical Safety Officer.

Maria Birkmyre RNA,SCPHN, Babylon Clinical Safety Officer

Maria has been working in clinical safety in health and social care software from 2008 onwards and has held CSO posts in several organisations. She has been an accredited clinician with NHS Digital for over 10 years .

Dr Saurabh Johri, Babylon Chief Scientist

Saurabh leads the AI Research team. Prior to Babylon, Saurabh spent 5 years as Post-doctoral research fellow at the MRC Centre for Outbreak Analysis and Modelling at Imperial College London, working in collaboration with the CDC and Gates Foundation.

This briefing covers four main elements:

1. A brief description of the different symptom checkers developed by Babylon, for whom each is made available, and the rationale for these decisions
2. Internal clinical safety assurance of the symptom checkers
3. Internal clinical safety assurance of broader clinical services provided by Babylon, including seeing a GP on smartphone or in person
4. External clinical safety assurance of all Babylon products and services, including the organisations involved in providing the external assurance.

Babylon's Symptom Checkers

Babylon provides a suite of digital health services, provided via mobile applications and on the web. One component of these services is the ability for people to use a chatbot to check their symptoms. Our experience has demonstrated that different users have different needs and priorities when it comes to checking symptoms. Some users value general information on possible causes for their symptoms, whilst for others signposting on what the right action to take ('triage') is most important. In parallel with developing a system that can safely and accurately provide both condition information and triage simultaneously, we have deployed two separate AI systems to cater for the different needs of our users.

- **Triage:** This system provides a user with advice regarding the most appropriate form of healthcare service to seek based on their medical symptoms, or where appropriate gives safe clinical advice (e.g., NHS Choices leaflets displayed on the smartphone). This system is built as curated triage flows, based on clinical evidence and acumen, where the questions and triage logic are managed by the Babylon clinical team, and undergoes thorough internal and external clinical safety testing before each release.
- **General information:** This system provides general information on conditions that may be possible causes of a user's symptoms. This is based on a probabilistic graphical model (PGM), an interpretable and auditable machine learning model which uses Bayesian inference to calculate the possible causes of the presented symptoms. This model has an understanding of links between symptoms, diseases and risk factors, using epidemiology data to understand the prevalence of conditions in the population. The system does not operate as a 'black box' and therefore each decision the system makes

can be explained. In addition, the model is not continuously learning in a live environment, meaning that the system can be fully tested for clinical safety, by internal and external clinicians, before each release. Clear labelling is provided throughout the user journey that this service “is only for general information purposes and isn’t a personal or medical diagnosis.”

Triage is an essential component of providing safe, appropriate care in the NHS, and is widely used across general practice, NHS111, 999 and A&Es. It is needed to make sure people are prioritised according to clinical need, and supports appropriate management of resources in a taxpayer-funded, free at the point of need service. For this reason, users of our ‘NHS Online: 111, powered by Babylon’ and the NHS ‘GP at Hand’ service have access to the Triage system and not to the system that provides General Information. The system they have access to therefore suggests the next step for the user to take – for example being given self-care advice when appropriate, advice to attend a local pharmacy, to see a GP, attend A&E, or call 999 when appropriate.

Internal Clinical Safety Assurance of the Symptom Checkers

Integral to our clinical safety approach is the clinical evaluation of all products, and the clinical safety and risk management processes that are in place across all products and services. To ensure the continued safety and effectiveness of any products in a live environment the above elements are regularly reviewed and updated, overseen by a Clinical Safety team, and documented in the Clinical Safety Management System, as per recognised NHS Digital safety standards.

The triage system described above is regulated under the EU Regulation on Medical Devices 2017/745, overseen by the MHRA in the UK, and is registered as a Class 1 Medical Device. The system is also CE marked, demonstrating compliance with all relevant European standards.

Prior to deployment of the triage system within the NHS service a thorough process of due diligence was undertaken alongside NHS England colleagues. A recent review of the hazard log - an industry-standard component of a clinical safety and risk management - conducted with representatives of NHS England and NHS Digital highlighted no significant concerns.

All triage algorithms undergo thorough clinical evaluation both pre- and post-market, with this process being repeated for all major updates since initial release. Pre-market testing aims to scrutinise the clinical safety of the product across both the product’s intended use and foreseeable misuse. This process involves external clinicians assessing the clinical safety of the product, across emergency, common, and ‘edge-case’ scenarios, as well as user testing being conducted to ensure appropriateness of language and clarity of

messaging. All identified clinical safety concerns have always been rectified before deployment into the live environment. All testing results since the product launch are documented within the Clinical Evaluation Report of the regulatory documentation.

Post-market surveillance and clinical governance procedures are critical to ensuring continued safety of the product in the live environment, as well as alerting us to any clinical safety concerns that have not become apparent despite rigorous pre-market testing. Post-market surveillance is undertaken in a number of ways, including regular auditing of clinical safety, user feedback, and complaints rectification and escalation. An example of the surveillance undertaken at Babylon is that following an update to the triage service two thousand user interactions with the service are clinically audited by GMC registered clinicians. All updates to the system, both routine and urgent are subject to a change management process which ensures the integrity of the change, and tests for any clinical safety regressions introduced by the change. All conversations with the service are logged, allowing full traceability of any feedback or complaints. A procedure is in place to comprehensively deal with all concerns or complaints in a timely and safe manner, and ensuring external escalation when appropriate. The product has been in use since 2016, and there have been no reported incidents of patient harm to date.

The General Information service described above has undergone a similar clinical safety and risk management process and will continue to be reviewed in line with routine post-market surveillance activities. Integral to the continuous improvement of this service is the feedback of Babylon GP's who not only are involved in pre-market testing, but also have visibility of recent chatbot conversations undertaken by their patients. The system does not act as a 'black box' and therefore there is explainability of each outcome provided by the service. In addition, each update is done in a controlled manner, preceded by formal pre-market testing, rather than undergoing continuous and autonomous updating of the live system. It has been live on the market since January 2018 with no reported incidents of patient harm.

Clinical governance procedures are in place to ensure that there is appropriate accountability and priority given to clinical safety both within the Artificial Intelligence teams and company-wide.

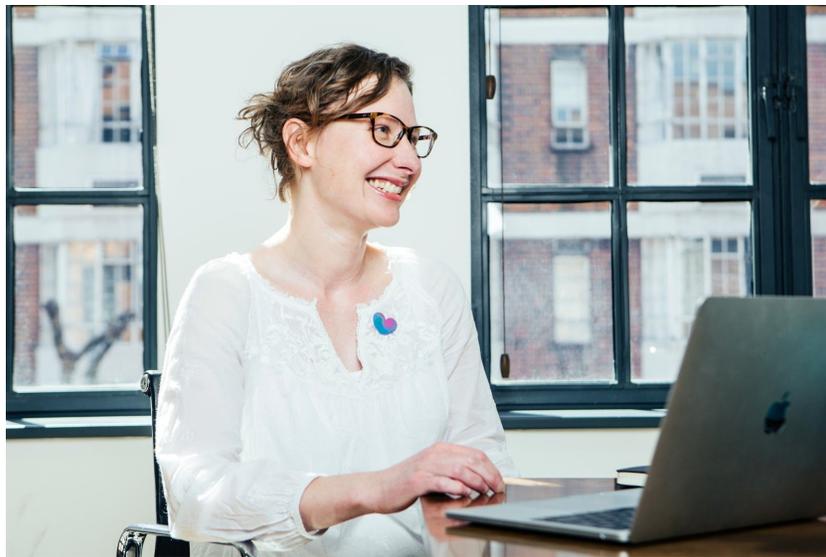
Internal Clinical Safety Assurance of Babylon Clinical Services

We set up Babylon not to replicate traditional primary care standards, but to combine clinical expertise with the power of technology to improve standards on safety, quality and accessibility. As we designed our service from scratch, we were able to set standards that were previously unheard of globally, and deploy these across both economically rich and poor nations.

Maintaining a positive culture

The 200 GPs who have already joined Babylon do so because they are treated with respect and not put under the levels of pressure all too common in today's pressured health care world. 95% of Babylon doctors surveyed say "Usually, I can manage the amount of my work well.". In comparison, the BMA showed in 2016 that 84% of GPs across the NHS reported that workload pressures are either "unmanageable" or "excessive" and having a direct impact on the quality and safety of the care they deliver to patients.

Every person working with us is encouraged to provide feedback on all elements of the babylon service.



The safest course of action for any patient with any concern is to be assessed in a timely manner to prevent conditions deteriorating or critical symptoms going unchecked. Our offer of a GP appointment within 2 hours, and booking in seconds, is in contrast to the current norm for routine appointments in many NHS practices of 1-2 weeks. As a result, people are much less likely to put off going to see the GP - a major factor in late diagnosis.

Our system of recording and archiving every video consultation for subsequent review, audit and learning is unparalleled in UK healthcare – and to the best of our knowledge worldwide. It resolves a serious safety concern in traditional general practice that recording currently relies exclusively on doctor's recall as transcribed through written notes. Our approach eliminates any possibility of disagreement between clinician and patient over what symptoms were discussed, what action was recommended, and provides both parties with an instant verbatim note and recording, introducing a new level of clarity and transparency into consultations. Patients have full, immediate access to the original consultation as well as the clinician's written notes.

Every quarter, we undertake a full clinical audit of recorded consultations. We have developed a bespoke, 20- point audit toolkit which is used to review a representative sample of every GP's consultations in detail. Combined with our real-time ratings system, which requires every patient to feed-back on their consultation, with all results published transparently, we have a very robust audit system down to the level of the individual clinician that surpasses standards in traditional primary care. To date over 85% of over 40,000 ratings we have been given for digital consultations are for the full five stars (our average star rating is 4.8 out of 5).

All ratings of 3 stars or less are investigated personally by the Senior Clinical Team with actions taken to remedy the issue tracked. Lessons learned are shared with the clinician community to facilitate learning and prevent reoccurrence. In addition, Babylon has put in place an overall safety system and culture that empowers front line staff in line with the "human factors" approach to safety improvement.

4. External assurance on Babylon products and services

Informal and formal external scrutiny and assurance is an important element of ensuring safety across health services, just as it is in any safety-critical industry.

Informal feedback

We welcome people's views, feedback and suggestions on our services - both positive and where improvements are necessary. People already use a variety of channels to contact us, from social media to emails to feedback provided through the babylon app. We are grateful for this and will continue to find ways to make it as easy as possible for people to provide this feedback. In particular, we are grateful to all GPs and other clinicians who have used their knowledge and experience to test the symptom checkers. All issues that are raised with us feed into the continuous improvement processes described above.

We are putting in place an easy process for people to identify any symptom checker outcome which they believe needs review. Put simply, we want as many people as possible to help us continuously improve so that an accessible and affordable health service for everyone on the planet can be made a reality as quickly as possible.

Formal assurance

The table below sets out the main formal external assurance processes that the different products and services powered by Babylon are subject to. Assurance processes specific to NHS services - 'GP at hand' and 'NHS Online: 111, powered by Babylon' - are shown separately, with the final column showing the processes that cover the Babylon products and services, independent of whether they are for use in the NHS or not.

In all cases, the results to date have been positive. The valuable scrutiny that external organisations bring helps us to learn and continually improve. We have never needed to withdraw a product or service from use.

	NHS - GP at hand	NHS Online: 111, powered by Babylon	Overall Babylon products and services
NHS England	<ul style="list-style-type: none"> ● Ongoing Clinical Review, led by NHS England Medical Directors in London ● Contractual compliance review ● Independent evaluation (by NHS England and Ipsos Mori) ● Assurance on DCB0129 and DCB0160 standards overseen by NHS England and NHS Digital 	<ul style="list-style-type: none"> ● Core Artificial Intelligence technology assured for “NHS 111 powered by Babylon” and “NHS Online:111”: <ul style="list-style-type: none"> ○ Clinical Reference Group sign-off ○ Outcomes assessment and testing from NHS 111 provider (LCW) ○ Commissioner and provider clinical leads ran 118 different historical patient scenarios through the Babylon software, the majority of which related to previous major or significant events. All the disposition outcomes were in line with NHS 	<ul style="list-style-type: none"> ● Clinical safety management plan details activities around DCB 0129 and DCB 0160 standards including delivery timelines.

		Pathways resulting in safe outcomes in each of the tests.	
NHS Digital	<ul style="list-style-type: none"> Assurance on DCB0129 and DCB0160 standards 	<ul style="list-style-type: none"> Assurance of data flows within and outside the N3 Network 	<ul style="list-style-type: none"> NHS Information Governance Toolkit NHS Digital Apps library assessment
Hammersmith & Fulham CCG	<ul style="list-style-type: none"> Ongoing assurance of contractual compliance Equalities Impact Assessment Equality and Diversity System self-assessment 		
MHRA	The symptom checker component of Babylon's service is subject to Medical Device regulation, which is overseen by the MHRA.		<ul style="list-style-type: none"> As per Class 1 Medical Device requirements the symptom checker meets all essential requirements, including clinical evaluation, technical documentation, and post-market surveillance procedures. Babylon Partners Limited MHRA registration Number 7139
CQC	<ul style="list-style-type: none"> Registration process 	<ul style="list-style-type: none"> n/a 	<ul style="list-style-type: none"> Registration process

	<ul style="list-style-type: none"> • Ongoing inspection and rating 		<ul style="list-style-type: none"> • Ongoing inspection and rating
Information Commissioner's Office	Subject to GDPR		<ul style="list-style-type: none"> • GDPR compliance
British Standards Institute			<ul style="list-style-type: none"> • Babylon Partners are certified as ISO27001 compliant