OpenSAFELY

Please describe the work that you wish to be considered for the John Perry Prize

OpenSAFELY is a secure, open analytics platform for NHS GP records. It enables researchers to analyse whole-population primary care data without moving the underlying records, combining unprecedented scale with privacy, transparency and reproducibility. It has transformed how routine GP data can be used rapidly for research, service improvement and safer care.

Please explain the relevance to primary care informatics of your submission

OpenSAFELY's is built on the foundation of primary care informatics. The platform makes electronic GP records routinely collected over decades available for research, surveillance and service improvement. GP records contain longitudinal information on consultations, diagnoses, prescriptions, tests, referrals and demographic characteristics. This enables primary care questions that cannot be answered reliably from small samples or extracted datasets: who is receiving care, where inequalities arise, how prescribing and clinical activity change, which treatments are effective and safe, and how outcomes vary between patient groups. The platform also supports a learning health system by providing reproducible analyses and updating dashboards. Its expansion beyond COVID-19 extends this capability across clinical conditions, making primary care data a practical component of national NHS analytics and research infrastructure.

Please explain the originality or innovation of your submission

OpenSAFELY addresses a problem at the heart of primary care informatics: how to use highly detailed, highly sensitive GP records at national scale without creating unacceptable privacy risks. Its innovation is architectural as well as analytical. Patient data remain in the secure environments where they are held; researchers develop analyses away from real patient records using dummy data, then submit code for automated execution against the real data. Results are aggregated and checked before release. The platform is also radically open: platform and analysis code are publicly shared, activity is publicly logged, and common data-preparation tools are standardised and reusable. Its federated design allows the same code to run across data centres without creating a central national copy of GP records. Together, these features combine privacy, transparency, reproducibility and scale in a way that previous extract-and-share approaches could not achieve

Please explain the impact or benefits of your submission

OpenSAFELY has already changed clinical and policy decisions at national scale. During COVID-19, its analyses of more than 17 million people rapidly identified risk factors that informed shielding and pandemic planning; work on learning disability and other risk groups contributed to JCVI vaccination priorities. Near-real-time monitoring identified inequalities in vaccine uptake and treatment access, supporting targeted NHS action. Real-world evaluations of COVID-19 treatments informed NICE guidance and international policy. The platform has also identified patient-safety problems: analysis found people with mechanical

heart valves receiving inappropriate anticoagulants, prompting a National Patient Safety Alert. More broadly, OpenSAFELY has enabled researchers to work across the NHS GP population, producing more than 100 published outputs and supporting over 200 projects across more than 30 organisations. Its impact is therefore not only individual studies, but a reusable infrastructure for better, safer and more equitable primary care.

Please explain the scalability of your submission

OpenSAFELY was designed for scale from the outset. Instead of creating and distributing copies of GP records, it takes analysis code to data held securely in existing environments. Standardised data-processing tools and an interoperable approach mean researchers can write an analysis once and run it across data centres, while avoiding a single centralised data lake and reducing the risk of vendor lock-in. The platform now supports analyses across more than 57 million patients' pseudonymised primary care records via the NHS England OpenSAFELY GP data service. We have deployed it in other data centres that contain health data and recently have embarked on projects to install OpenSAFELY internationally. The same architectural principles are already being applied beyond healthcare, including school data, demonstrating transferability beyond the original use case. Because the platform, tools, documentation and analysis code are openly shared, other organisations can inspect, reuse and adapt them. This makes OpenSAFELY a scalable model for trustworthy population-level analytics rather than a bespoke project tied to one dataset or institution.

Please describe the evidence base of your submission

The evidence base combines independent evaluation, large-scale research outputs and real-world operational use. OpenSAFELY's first major study, published in Nature, analysed more than 17 million people within 42 days of the project beginning and identified major COVID-19 risk factors. The platform subsequently supported peer-reviewed studies in leading journals, live vaccination and treatment monitoring, and analyses used in national clinical and policy decisions, including NICE guidance and a National Patient Safety Alert. By spring 2026, OpenSAFELY had supported more than 200 projects, over 100 published outputs and users across more than 30 organisations, with its work receiving more than 10,000 citations. Trustworthiness has also been tested: citizens' juries commissioned by the NHS, National Data Guardian and NIHR gave OpenSAFELY the strongest support among the data initiatives reviewed, with 77% very much in support and 87% supporting continuation beyond COVID-19. The platform has also received the MRC Open Science Impact Award and CogX award for open-source innovation. And the prestigious Queen Elizabeth Prize from the King. More information about OpenSAFELY is available at https://www.opensafely.org/