

## Introduction to the IMI2 Big Data for Better Outcomes Programme (BD4BO)

**All information regarding future IMI Call topics is indicative and subject to change. Final information about future IMI Calls will be communicated after approval by the IMI Governing Board.**

The IMI2 Big Data for Better Outcomes (BD4BO) programme **aims to catalyse and support the evolution towards value-based and more outcomes-focused sustainable and therefore better quality healthcare systems in Europe, exploiting the opportunities offered by the wealth of emerging data from many evolving data sources** by generating methodologies and data that will inform policy debates. The programme's objectives are to maximise the potential of large amounts of data from variable, quickly-developing digital and non-digital sources which will be referred to as 'big data' in the context of this initiative.

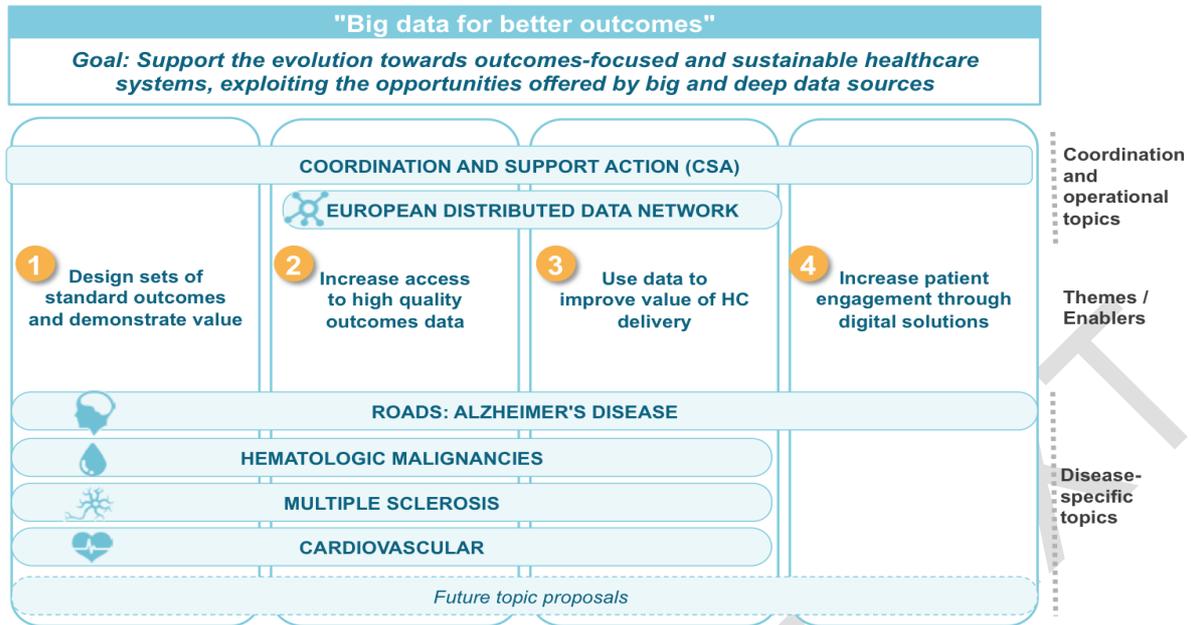
This programme will provide a platform and resources for defining and developing enablers of the outcomes transparency evolution together with patients, payers, physicians, regulators, academic researchers, healthcare decision makers, etc. The key enablers are:

- definition of outcome metrics
- protocols, processes and tools to access high quality data
- methodologies and analytics to drive improvements
- digital and other solutions that increase patient engagement.

### Programme structure

The programme will be composed of several topics which will address key enablers for the transition of healthcare systems towards more outcomes transparency, including an over-arching coordination structure (through a Coordination and Support Actions (CSA)), key structural and technology components (European Distributed Data Network) and several disease/therapeutic area (TA) topics focusing on a specific disease, population, therapeutic area or technology. Only one proposal under each topic will be selected.

In this Call, the BD4BO Distributed Data Network is launched.



**Figure 1: Programme structure, themes / enablers and CSA**

The success of the overall programme will rely on a coordinated approach across projects to ensure strategic alignment and consistency and to define new business and health funding models (including incentive models) that will allow for healthcare systems transformation. In addition, integration of areas of expertise which are common to most projects (such as legal, ethics, data privacy, sustainability or collaboration with payers/HTAs) will yield higher quality results, consistency and increased efficiency by avoiding duplication of work.

Two projects, the Coordination and Support Actions (CSA), and this call for a European Distributed Data Network Project (DDN) will therefore offer services to and complement activities of disease/therapeutic area related projects through:

- a central repository of knowledge/information;
- a common ethical and personal data protection review and advice;
- common standards for the collection, analysis and management of personal level data/knowledge;
- assistance on the implementation of common data models and in the aggregation of data from different sources.

The distribution of tasks with responsibilities across different project teams within the programme (subject to adjustments as projects evolve) is summarised in figure 2:

	CSA	DDN	TA projects
Consistency & quality	<ul style="list-style-type: none"> <li>• policy/direction, objectives, call texts</li> <li>• facilitate interactions and learnings between projects</li> </ul>	<ul style="list-style-type: none"> <li>• Standards for data and knowledge</li> </ul>	<ul style="list-style-type: none"> <li>• Contribute to Data Standards</li> <li>• Define TA specific Data Standards</li> <li>• Implementation</li> </ul>
Knowledge integration	<ul style="list-style-type: none"> <li>• Policy principles (programme level)</li> <li>• Gap analysis and recommendations for projects</li> </ul>		<ul style="list-style-type: none"> <li>• TA specific recommendations</li> </ul>
Knowledge repository/platform	<ul style="list-style-type: none"> <li>• Processes for collection, use and exploitation of knowledge and active process management</li> </ul>	<ul style="list-style-type: none"> <li>• Repository infrastructure</li> </ul>	<ul style="list-style-type: none"> <li>• Generate and manage TA specific contents</li> </ul>
Engagement with HC stakeholders & communication	<ul style="list-style-type: none"> <li>• Branding (templates, guidance)</li> <li>• Programme comms material</li> <li>• Programme website</li> <li>• Stakeholders engagement</li> </ul>	<ul style="list-style-type: none"> <li>• Recommendations on data formats and standards</li> </ul>	<ul style="list-style-type: none"> <li>• TA specific interactions with stakeholders (+reporting to CSA)</li> <li>• Communication (project level)</li> </ul>
Data protection & integrity	<ul style="list-style-type: none"> <li>• Legal &amp; ethical standards</li> <li>• Code of practice</li> <li>• Common ethics Board</li> <li>• Templates/guidance</li> </ul>	<ul style="list-style-type: none"> <li>• Technical data security and privacy solutions</li> <li>• Common minimum data standards</li> </ul>	<ul style="list-style-type: none"> <li>• Operational interactions with competent authorities</li> </ul>
Data sustainability & growth of networks		<ul style="list-style-type: none"> <li>• Data sustainability mechanisms (including quality of data and adequate use)</li> </ul>	

**Figure 2: Allocation of tasks between Coordination & Support Action, Distributed Data Network and therapeutic area focused projects**

### Current status of the BD4BO programme

So far in this programme, 2 calls for proposals have been launched in 2015 (Calls 6 and 7). The winning short and full proposals will, if evaluated positively, start the project in 2016.

- Real World Outcomes Across the AD (Alzheimer’s disease) Spectrum (ROADS) to Better Care (indicative budget: EUR 4 000 000 + 4 000 000);
- development of an outcomes-focused platform to empower policy makers and clinicians to optimise care for patients with haematologic malignancies (indicative budget: EUR 20 000 000 + 20 000 000);
- coordination and support action (CSA) for the big data for better outcomes programme (indicative budget: EUR 3 550 000 + 3 550 000);
- increase access and use of high quality data to improve clinical outcomes in heart failure (HF), atrial fibrillation (AF), and acute coronary syndrome (ACS) patients (indicative budget: EUR 9 672 000 + 9 672 000).

These 4 call topics represent a total public-private investment of EUR 74 444 000.

### Collaboration agreements

To ensure the interactions between the projects under the BD4BO programme, the therapeutic area/disease (TA) projects are expected to actively contribute key horizontal results to the Coordination and Support Actions (CSA) and the current call for a European Distributed Data Network (DDN) projects which will provide direct advice and support to the TA projects. Therefore all grants awarded for the TA projects will be complementary to the Grant Agreements under the CSA and DDN topics. The respective options of Article 2, Article 31.6 and Article 41.4 of the Model Grant Agreement will be applied.

The TA consortia will conclude collaboration agreements with the Coordination and Support Actions consortium and the current call for a European Distributed Data Network consortium. The collaboration

agreements are expected to include details of the services provided by the CSA and DDN to the TA specific projects such as the provision of data collection standards and processes, an interim repository for knowledge storage and management, data privacy standards, compliance and ethics regulations, including templates and other operational support.

The TA-specific projects are expected to contribute to the CSA knowledge repository and integration of learnings, and also participate in joint advisory boards and coordination boards to align on strategic programme elements such as definition of health outcome measurements, operational standards including data and knowledge collection and aggregation standards, common usage of IT infrastructures, communication of programme results and operational issues indicated in figure 2. All TA projects should ring-fence resources for these activities (approximately 5% on average, for example, experts to participate in central programme boards, participate in the adoption, adaptation and/or definition of common data standards, and/or cash that will cover the cost of operationalising e.g. central ethical and data protection boards and maintenance of the common IT infrastructure).

## Need and opportunity for public-private collaborative research

The Big Data for Better Outcomes programme aims to provide high quality information that may provide decision makers with evidence on the enablers of value-based healthcare systems focusing on health outcomes. This healthcare system transformation would encompass payments, consider value and support aligned incentives between primary and secondary care moving towards the common goal of superior healthcare delivery and high quality data being made available. Therefore the engagement of patient organisations, regulators, payers, providers and other public stakeholders throughout the BD4BO programme is essential to ensure findings from those projects have appropriate buy-in and ultimately deliver real impact in transforming healthcare systems.

### .Expected impact

The expected impact of the programme would be a comprehensive plan for the development and implementation of key enablers to support the evolution towards value-based and more outcomes-focused and sustainable healthcare systems in Europe, exploiting the opportunities offered by big and deep data sources. The programme will also enable evolution and management of R&D portfolios and prioritisation of research methodologies in line with an outcomes focus.

Applicants should refer to the 'expected impact' sections under each of the BD4BO topics.

## Topic: A European distributed data network (DDN) project to facilitate intra-EU access to, and analysis of, Real World Data to improve health outcomes for EU patients

### Big Data for Better Outcomes Programme (BD4BO)

**All information regarding future IMI Call topics is indicative and subject to change. Final information about future IMI Calls will be communicated after approval by the IMI Governing Board.**

#### Topic details

Action type	Research and Innovation Actions (RIA)
Submission & evaluation process	2 Stages

**Please Note:** A Work Package (WP) focused on education and training is foreseen to be included in this topic. This WP is dependent on the participation of an associated partner and their in kind contribution. At the time of launching the consultation the association of this organisation to IMI2 has not been finalized and should it not be completed prior to call launch the intention would be to remove the education and training WP. Removal of this work package would not endanger the core activities of the topic and not prevent the resultant project from achieving its key objectives. In such an instance alternative routes would be sought to take these education and training activities forward either as part of another topic or as a topic in its own right under the BD4BO programme.

All activities related to the education and training WP in the current topic text are identified by text enclosed in square brackets.

#### Specific challenges to be addressed

Europe has a variety of electronic sources of real world data and a long history of investing in health informatics. Although there is potential to improve health outcomes by leveraging on existing data, accessing this data for research purposes has been hampered by regional fragmentation, lack of interoperability and privacy concerns.

Current technology offers opportunities to capture vast amounts of data, with different scope and characteristics. Claims data provide comprehensive information on the patient's interaction with the health system, including diagnoses or drug prescription filling. While not as rich as electronic health record data, it is extremely broad and granular with respect to treatment patterns. Electronic health records (EHR) provide rich information on the actual health status of the patient, capturing data such as lab results, vital signs, habits or non-prescription drugs. Moreover, patient-reported outcomes (PROs) or selected demographic data can fill in gaps for information not covered by the previous sources.

The combination of different data sources can unlock powerful insights that can be used to improve health outcomes in several ways. The access to quality real world data is needed to drive improvements in the delivery of care, assess the effects of products in the market, identify unmet needs or opportunities to develop new products and increase patient engagement.

However, access to and combination of data sources faces significant obstacles. First, fragmentation and lack of linkages among data sources prevents effective aggregation of diverse data. Second, lack of agreement on common standards hampers the interoperability and comparability of data. Third, privacy and regulatory concerns hinder the integration of data from different sources. Surmounting these obstacles to facilitate pan-EU querying of data is a necessary foundational step to advancing in the use of big data to support health outcomes analysis. Fortunately much work has already been done through existing IMI projects (e.g. EMIF) which will serve as a foundation and a model for what is envisaged in the proposed project.

*[Europe also lacks today a dedicated education and training infrastructure and contents to train next generation scientists specialised in use and analysis of big data in biomedical research, particularly:*

- *big data generation and processing aligned with regulatory and increasingly adaptive models which connect high number of data points and a variety of data sources;*
- *use of big data in decision making through healthcare and healthcare products pathways].*

## Need and opportunity for public-private collaborative research

The development of a distributed platform, where patient-level data would remain hosted in the country/organization of origin but queried remotely by a central coordination server, would enable the access to data, while addressing some of the key barriers currently hampering it.

Accessing data from different participating entities requires taking into account legal requirements and local restrictions. A public-private approach will be necessary to facilitate cross-border cooperation and provide assurances about safeguarding of data and patient privacy. Developing the distributed data network will also require alignment on the research questions to be answered and adoption of a limited number of common data models, which will necessitate a broad consortium approach.

IMI projects (such as EMIF<sup>1</sup> and EHR4CR<sup>2</sup>) have proven that effective collaboration between public and private entities is required to overcome some of the barriers preventing access to data from different sources. There is significant potential to build on these efforts, while expanding the scope by involving the necessary public stakeholders (payers, HTAs, regulators, patient advocacy groups) to build trust and provide confidence to data providers.

*[Setting a European programme and a centre of excellence for education and training is also best delivered in public private collaboration - combining the input on the needs and standards required by regulators and industry processes and sustainable education infrastructure anchored in the academic system that can seamlessly integrate in higher education and continued professional development schemes.]*

Both public and private parties need to have a seat at the table in the development of this initiative, and they can also contribute different kinds of expertise. Private entities have expertise in data capture and they are the owners of a considerable part of the data, while benefitting from the potential combination with other datasets to generate insights. Public entities are also data *[and higher educational infrastructure]* owners, they can mobilize other parties and they can bring the trust that is required to establish a governance framework.

## Scope

---

<sup>1</sup> [www.emif.eu](http://www.emif.eu)

<sup>2</sup> [www.ehr4cr.eu](http://www.ehr4cr.eu)

This call topic is specifically to

- scope out the scale and resources needed to upscale a distributed network from 20 data sources to 250 data sources;
- scope out the opportunity afforded by integrating emerging data types such as patient derived data;
- scope out the opportunity afforded by expansion of the use of these types of data to new stakeholder groups such as physicians or even patients;
- provide support for ongoing BD4BO disease specific projects (e.g., bridging data source access and analytic platform infrastructure);
- assistance for those projects in loading data into the EMIF platform;
- *[develop an Education and Training centre or network of excellence].*

## Expected key deliverables

- Clear proposal/project plan for scaling EMIF from ~20 data sources to 250 by 2025, enabling inclusion of novel data types and use of the data by new stakeholder groups;
- definition and clarification of value proposition for data providers including an agreed strategy that will maximise the number of data providers that are willing to put data into the DDN;
- disease specific BD4BO projects have resource and capability to enable access to the multiple data repositories that these projects identify, including guidance on mapping to the common data model;
- support for the overall BD4BO CSA knowledge repository;
- *[Education and Training Cursum plan and contents against identified needs and rapidly evolving technology and regulatory environment, including sustainability and upscaling plan].*

## Expected impact

The expected impact of this project can be summarised into the following areas:

- a significantly increased access to Real World Data (RWD) in terms of diversity of data types, number of data providers accessible and breadth of stakeholders that can access the data;
- to pave the way for a distributed network built upon 250 data sources by 2025;
- increased visibility of legal and privacy constraints across data providers ensuring compliant access to data is ensured;
- extension of several related IMI initiatives and with the greater BD4BO programme creation of a co-ordinated RWD ecosystem across Europe;
- ultimately it is expected that better more controlled access to data will result in improved understanding of disease and its progression, at the individual patient and population level, leading to improved health outcomes across a whole variety of diseases;
- *[Development of expertise in utilisation of big data in healthcare products development and healthcare pathways taking account of European regulatory context and evolving technology will speed up uptake of new technologies in R&D and in decision making thereby contributing to modernising both healthcare systems and healthcare produces development].*

## Potential synergies with existing consortia

Synergies encompass both technical and policy-related issues currently being advanced by other current and prior IMI initiatives. For instance, technical includes building upon EMIF efforts pertaining to data models and infrastructure for broad access and analysis. Policy-related aspects including models for safeguarding data privacy, facilitating data sharing arrangements, governance for vetting research questions.

The goal of a distributed data network approach is clearly aligned closely with other IMI initiatives, such as EMIF, EHR4CR, GetReal, and PROTECT, all of which are concerned with facilitating analysis of real world data and exploring impact of real world outcomes. Close collaboration with EMIF will be required to build on the existing work and expand its scope. Other emerging IMI projects such as those under the Remote Assessment of Disease and Relapse Programme (for actigraphy type data) could also become important as those projects mature.

Collaboration strategies with similar non-IMI initiatives such as Clinical Practice Research Datalink (CPRD), aimed at linking anonymised clinical data from the National Health Service in the UK for observational research, will also be analysed.

Moreover, this initiative will coordinate with the Big Data for Better Outcomes CSA and support the objectives of other projects under the programme which address the theme of access to quality outcomes data. For instance, it is anticipated that the DDN platform can serve as a source of data as well as a potential knowledge repository for other projects, should there be a need. Also, by broadening the scope and range of stakeholders contributing and benefiting from the network, there is increased opportunity to drive sustainability over time.

*[For the purpose of Education and Training component, the consortium will synergise with existing IMI Pharmatrain and other complementary initiatives still to be mapped.]*

## Indicative duration of the project

The indicative duration of this action is 24 months.

### Future project expansion

Potential applicants must be aware that the Innovative Medicines Initiative 2 (IMI2) Joint Undertaking may, if exceptionally needed, publish at a later stage another Call for proposals restricted to the consortium already selected under this topic, in order to enhance their results and achievements by extending their duration and funding. The consortium will be entitled to open to other beneficiaries as they see fit.

Such further work would be the natural progression of the project leveraging any success achieved. Building on these prior successes and positive results would maximise the long term impact of the larger project. Any proposed project extension would also take advantage of already established collaborations and networks forged in the overall project, thereby maximising efficiency on time and resources. A restricted Call would achieve this in the most efficient way. The detailed scope of the restricted Call shall be described in the relevant Annual Work Plan.

## Applicant consortium

The applicant consortium will be selected on the basis of their submitted short proposals.

The applicant consortium is expected to address all the research objectives and make key contributions to the defined deliverables in synergy with the industry consortium which will join the selected applicant consortium in preparation of the full proposal for stage 2.

Specifically the applicant consortium will be expected to formulate a detailed strategy based upon the experience of the ongoing EMIF project scaling up to encompass up to 250 data sources by 2025. This will include elements that describe solutions to the technical and practical challenges that an ambitious goal such as this is likely to present. In addition, the consortium will need to work closely with the other BD4BO projects to ensure that their data can be effectively accessed via, for instance, the current EMIF infrastructure. A key component of this will be bridging access to applicable data resources and infrastructure for ongoing disease-specific BD4BO projects from EMIF to Phase 1 to Phase 2.

*[Further, the applicant consortium expects to include public higher education institutions with a crucial proficiency in transferring knowledge and technology linked to big data collection, visualisation and analysis. In addition, this includes training and tutelage in developing expertise in definition and implementation of a common data model, as well as in the aggregation of data from different sources].*

## Suggested architecture of the full proposal

The applicant should include in their short proposal their suggestions for creating the full proposal architecture, taking into consideration the industry contributions and expertise as indicated.

The final architecture of the full proposal will be defined together with the industry consortium and should enable activities designed to achieve all objectives and deliverables as indicated in the previous relevant sections and in collaboration with the EFPIA and associated partners. The final architecture of the full proposal will be defined by the participants in observance of IMI2 rules and in contemplation of the achievement of the project objectives.

In the spirit of the partnership, and to reflect that IMI2 Call topics are built upon identified scientific priorities agreed together with EFPIA beneficiaries/large industrial beneficiaries, it is envisaged that IMI2 proposals and projects may allocate a leading role within the consortium to an EFPIA beneficiary/large industrial beneficiary. Within an applicant consortium discussing the full proposal to be submitted at stage 2, it is expected that one of the EFPIA beneficiaries/large industrial beneficiaries may elect to become the coordinator or the project leader. Therefore to facilitate the formation of the final consortium, all beneficiaries are encouraged to discuss the weighting of responsibilities and priorities therein. Until the roles are formally appointed through a consortium agreement the proposed project leader shall facilitate an efficient negotiation of project content and required agreements.

## Glossary

BD4BO	Big Data for Better Outcomes
CPRD	Clinical Practice Research Datalink
CSA	Coordination and Support Action
DDN	Distributed Data Network

EHR	Electronic Health Records
EHR4CR	Electronic Health Records for Clinical Trials
EMIF	European Medical Information Framework
OMOP	Observational Medical Outcomes Partnership
PROs	Patient-reported Outcomes

INDICATIVE TEXT