



**Abschlussbericht des Forschungsvorhabens: „Europäische Initiative zum Aufbau eines Datenverbunds aus Ernährungsstudien mit Schwerpunkt Phänotypisierung (ENPADASI)“**

Zahlungsempfänger: Technische Universität München

Förderkennzeichen: 2814ERA07F

Vorhabenbezeichnung: Initiative zum Aufbau eines Datenverbunds aus Ernährungsstudien mit Schwerpunkt Phänotypisierung (ENPADASI)“

Laufzeit des Vorhabens: 12. Dezember 2014 bis 30. Juni 2017

Berichtszeitraum: 12. Dezember 2014 bis 30.06.2017

## **Beitrag des Ergebnisses zu den förderpolitischen Zielen, auch zwecks Evaluierung von Förderprogrammen**

Die Joint Programming-Initiative (JPI) wurde als Instrument zur besseren Harmonisierung und Koordination von Forschungsinitiativen in der EU (und externen Partnern) begründet. Da mehr als 85% öffentlich finanzierter Forschung noch immer auf nationaler Ebene finanziert, koordiniert, reguliert und evaluiert wird (*Current practices and experiences in the area of food and health research. EU-publication, published 2012-05-29. ISBN 978-92-79-21791-3*), erscheint die Kooperation und Koordination auf transnationaler Ebene sinnvoll. Ziele der JPI sind somit u.a.:

- die Koordination thematischer Schwerpunkte von Forschungsprogrammen in Europa und die Vermeidung von Redundanzen;
- die Förderung wissenschaftlicher Exzellenz durch gemeinsame Aktivitäten mit gemeinsamer Finanzierung, mit den Zielen einer Minimierung der Forschungsfragmentierung sowie einer effizienten und transparenten Nutzung öffentlicher Ressourcen;
- die Unterstützung grenzüberschreitender Kooperationen, die Förderung standardisierter Datenerfassung und die Erleichterung des Datenaustauschs.

ENPADASI liefert somit einen Beitrag zur Harmonisierung der Wissenschaft und zur Transformation der Befunde in die Gesellschaft und damit letztlich zur Förderung von Wohlbefinden und Gesundheit der Bevölkerung. Durch die Erstellung einer open-access Datenbank mit standardisierter Datensammlung und Datenspeicherung wurde dazu ein wichtiger Schritt getan. Mit Beteiligung von Partnern aus verschiedenen europäischen Ländern (siehe: <http://www.enpadasi.eu/partners.html>) ließ sich die Standardisierung und Vereinheitlichung über eine Vielzahl von national-geförderten Studien herstellen. Gleichzeitig führte die Standardisierung dazu, den Datenaustausch zu erleichtern und den Vergleich von Studien zu ermöglichen. Mit der Unterzeichnung eines *Data Transfer Agreements* durch die Partner wurde dies auch formalisiert. Die integrierte Datenanalyse lässt somit die Erhöhung der Fallzahl in vergleichbaren Studien zu und ermöglicht besser Meta-Analysen zu einer gegebenen wissenschaftlichen Fragestellung. In diesen Möglichkeiten liegt auch der besondere Wert mit entsprechendem gesellschaftlichen und wirtschaftlichen Nutzen von ENPADASI.

## **Wissenschaftlich-technische Ergebnisse des Vorhabens im Vergleich zu den ursprünglichen Zielen, erreichte Nebenergebnisse und gesammelte wesentliche Erfahrungen**

Das Ziel von ENPADASI, eine *open-access*-basierte Forschungsinfrastruktur zu schaffen, wurde mittels der Entwicklung einer Datenbank und eines dazugehörigen Analysensystems erreicht (DASH-IN; Abkürzung für: Data Sharing Initiative for Nutrition). Die Datenbank enthält zum Zeitpunkt der Erstellung dieses Abschlussberichts 79 Studien (Deliverable 2.1.1). Diese Datenbank ist eine grundlegend überarbeitete, erheblich erweiterte und benutzerfreundlich gestaltete Version der ursprünglichen sog. *Phenotype-Datenbank* ([www.dbnp.org](http://www.dbnp.org)). Die Integration neuer Studien führte zur Identifizierung und Behebung von vorhandenen Datenbankfehlern. Weiterhin wurde die Datenbank um neue technische Möglichkeiten (u.a. für den Import von Daten und der Konversion von Formaten) erweitert. Dies war Aufgabe von WP3, an dem TUM nicht beteiligt war. Eine technische Definition der Datenbankinfrastruktur, die Interventions- und Observationsstudien verbindet, ist eine besondere Herausforderung. Basierend auf den vorhandenen Studien wurde dazu eine einheitliche Terminologie entwickelt und mit existierenden Ontologien verglichen. Die Partner in ENPADASI haben dazu in Arbeitstreffen unter Beteiligung der TUM festgelegt, welche Einzelheiten in der Datenbank gespeichert werden sollen (Deliverable 2.2.1) und unter welchen Umständen der Austausch von Daten möglich wird (Deliverables 2.5.1 und 5.4). Die Einzelheiten zu den Deliverables stehen unter <http://www.enpadasi.eu/deliverables.html> zur Verfügung.

Das Web-basierte Analysesystem von DASH-IN (Deliverable 3.7.1) erlaubt die Integration der Studien in einen formatierten Rahmen. Die DASH-IN Infrastruktur verbindet die Daten-Server verschiedener Partnern des ENPADASI Projekts (nicht TUM). Diese Partner haben zugesagt, diese Server mit Ende des Projekts für weitere fünf Jahre auf eigene Kosten in Betrieb zu halten. Somit ist Fortbestand der Infrastruktur für mindestens fünf Jahre gesichert. DASH-IN erlaubt dabei grundsätzlich auch eine lokale Datenspeicherung, soweit dies im Einklang steht mit lokalen Datenschutzbestimmungen. Weiterhin wird eine zentrale Datenspeicherung von TNO (Niederländische Organisation für angewandte naturwissenschaftliche Forschung: *Toegepast Natuurwetenschappelijk Onderzoek*) vorgehalten. Metadaten können

dabei auch in der Datenbank von DASH-IN oder in Mica gespeichert werden. Mica ist Internet-basierte datenbankübergreifende Applikation für die Verwaltung und Veröffentlichung von epidemiologischen Daten (siehe auch: <https://www.obiba.org/pages/products/mica/>). Die *Phenotype*-Datenbank folgt den FAIR Prinzipien. FAIR steht für Findability, Accessibility, Interoperability, and Reuse of digital assets. Die Daten in der Datenbank müssen also auffindbar, zugänglich, interoperabel und wiederverwendbar sein.

Ergänzend und unterstützend zur Datenbank und deren Verwendung sind im Projekt ENPADASI entwickelt worden:

- eine Sammelstelle von *online* Quellen bzw. Hilfsmitteln zu ethischen und datenschutzrechtlichen Aspekten sowie zur Regelung und Handhabung von geistigem Eigentum (*Intellectual Property, IP*).
- eine Sammlung von Standardarbeitsanweisungen (*Standard Operating Procedures, SOPs*) für die Datenerhebung
- Schulungsmöglichkeiten für die Nutzung der Datenbank und des Analysesystems
- ein Dokument mit häufig gestellten Fragen (FAQ) zur Nutzung der Datenbank.

ENPADASI dient als Initiative der Förderung des Austauschs von Forschungsdaten. Im Projektverlauf wurden jedoch auch eine Reihe ethischer und rechtlicher Probleme identifiziert, die diesen Austausch behindern können. Daher wurden zunächst nur Metadaten und keine Rohdaten verfügbar gemacht. Mit der Unterzeichnung einer Vereinbarung zur Datenweitergabe (*Data Transfer Agreement, DTA*) wird dann den Wissenschaftlern – vorbehaltlich einer Einigung der Partner – auch die Bereitstellung der Rohdaten ermöglicht. Um Probleme beim Datenaustausch zu verhindern, wird aus Sicht von ENPADASI langfristig eine geeignete Kommission für die Regelung des Datenaustausches zuständig sein müssen.

Der Lehrstuhl für Ernährungsphysiologie der Technischen Universität München hat als ENPADASI-Partner die Daten zweier Interventionsstudien mit allen Befunden zur Verfügung gestellt und in die Datenbank eingebunden. Dabei handelte es sich um folgende Studien:

1. MIPROMET: Die Wirkung verschiedener Milchfraktionen auf die metabolischen Antworten bei gesunden und prädiabetischen Probanden;

2. HUMET: Die dynamische Variation der physiologischen Antwort auf verschiedene metabolomische Herausforderungen (Bsp. 36h fasten, oraler Glucosebelastungstest u.a. mit jeweiligen Zeitprofilen und mit einem Datensatz von mehreren Hundert Metaboliten in über 900 Plasmaproben).

Zusätzlich war der Lehrstuhl in enger Zusammenarbeit mit anderen Partnern an der Bearbeitung folgender Aufgaben beteiligt:

- Definition einer Ontologie für die Beschreibung von Ernährungsstudien
- Festlegung eines Katalogs von Minimalanforderungen für die Eignung von Daten zur Einbindung in ENPADASI
- Feststellung von Prüfkriterien zur Erfüllung jener Minimalanforderungen
- Entwicklung eines Bewertungssystems für die wissenschaftliche Qualität von Daten
- Beurteilung der Qualität der vorhandenen Daten.

Die erstellte Ontologie für die Beschreibung von Ernährungsstudien steht zur Verfügung unter <https://github.com/FrancescoVit/Ontology-for-Nutritional-Studies>. Die Minimalanforderungen für die Eignung von Studiendaten sowie die Prüfkriterien werden beschrieben im Anhang von Deliverable 2.3.1 siehe: [http://www.enpadasi.eu/deliverables\\_final\\_pdf/D2.3.1.pdf](http://www.enpadasi.eu/deliverables_final_pdf/D2.3.1.pdf).

### **Darstellung und Erläuterung der Angemessenheit von Aufwand und Zeit**

Die verfügbare Zeit erlaubte die Erfüllung aller geplanter Aufgaben. Das Verhältnis von Aufwand und Zeit zum Ertrag kann als angemessen klassifiziert werden. TUM war hauptsächlich in Arbeitspaket 2 (WP2) eingebunden. Für die Zusammenstellung und Bereitstellung der Studiendaten sowie die Eintragung in der Datenbank (Aufgabe 2.2) waren 18 Monate vorgesehen und dies konnte zeitgerecht realisiert werden. TUM war weiterhin bei der Erstellung von Minimalanforderungen (Aufgabe 2.3) und der Evaluierung von Studiendaten (Aufgabe 2.4) involviert wofür 18 bzw. 12 Monate zur Verfügung standen. Diese Aufgaben konnten dank der Laufzeitverlängerung ebenso erfolgreich abgeschlossen werden. Die Aufgaben in Arbeitspaket 1 (WP1), in dem es um die allgemeine Projektverwaltung handelte, sowie die Aufgaben 2.3 bis 2.5 wurden zum großen Teil in gemeinsamen Arbeitsgesprächen erfüllt, für die mit wiederholten Treffen ein Zeitraum von 24 Monaten vorgesehen war.

## **Arbeiten die zu keiner Lösung geführt haben**

Alle Aufgaben, an denen die TUM beteiligt, konnten einer Lösung zugeführt werden. In Bezug auf das Gesamt-Projekt ENPADASI war ursprünglich eine mehrstufige Datenanalyse vorgesehen. Es wurde jedoch deutlich, dass dafür schon einige Teilansätze in anderen Projekten erarbeitet waren und daher alternativ und kostenneutral eine Übersicht aktueller und verfügbarer Hilfsmittel für die Datenanalyse im Arbeitspaket erstellt.

## **Darstellung und Erläuterung der wissenschaftlichen und ggf.**

### **Anschlussfähigkeit für eine mögliche nächste Phase**

Eine komplette Unterstützung und Instandhaltung des Systems (Deliverables 1.2.3/4) benötigt jährlich ca. 100.000 – 500.000 Euro. Die Höhe dieses Betrags hängt ab von der Instandhaltungsebene des Systems und dem Umfang an neuen Daten, die in die Datenbank eingestellt werden. Sie beinhaltet Serverkosten, Personalkosten für Programmierer und Betriebskosten für die Nutzung des Systems. Eine Minimalbetrieb bedeutet noch immer Serverkosten sowie die Kosten für eine online-Unterstützung und für die Behebung von kritischen Fehlern. Ein hoher Verfügbarkeitsgrad beinhaltet zusätzlich auch die Weiterentwicklung der Infrastruktur sowie „Vor-Ort-Workshops“ zur Schulung in Anwendungen.

Mit der beachtlichen Anfangsinvestition und der Bereitschaft der Forschungsgemeinschaft zum Datenaustausch ist ENPADASI als funktionsfähiges System mit Strahlkraft auch über europäische Grenzen hinaus etabliert worden. ENPADASI weiterzuführen und weiterhin zu finanzieren muss daher ein primäres Anliegen sein. Für den weiteren Erfolg der Datenbank ist aber auch die beständige Aktualisierung zu ethischen und datenschutzrechtlichen Aspekte relevant. Dies könnte zukünftig beispielsweise in Zusammenarbeit mit CORBEL (Coordinated Research Infrastructures Building Enduring Life-Science Services) durchgeführt werden.

Auch die Aktualität der Standardprotokolle (SOPs) für die Datenerhebung muss sichergestellt sein, da sie den technischen Entwicklungen in der ernährungswissenschaftlichen Forschung folgen. Die Eintragung neuer Daten in das System

führt zwangsläufig auch zur Notwendigkeit für die Erstellung neuer Ontologie-Begriffe und somit der Erweiterung der ONS (*Ontology of Nutritional Studies*).

Hoffnung auf eine Weiterfinanzierung der Initiative besteht u.a. im Bündnis des Projekts ENPADASI mit ELIXIR (siehe: <https://www.elixir-europe.org/>) als einer Europäischen Infrastruktur für biologische Information zur Unterstützung der *Life Sciences* und zur Translation in Medizin, Agrarwissenschaften und Bioindustrie in Europa. Eine mögliche weitere Finanzierung könnte auch über den anstehenden Horizon 2020 „Foodcloud“-Call (Kennung: *DT-SFS-26-2019*; unter: [https://ec.europa.eu/commission/index\\_en](https://ec.europa.eu/commission/index_en)) erreicht werden. ENPADASI ist darüber hinaus assoziiert mit den Forschungsinfrastrukturinitiativen von FNH (Food, Nutrition and Health). Das Hauptziel dabei ist es, Befunde aus Lebensmittel-, Ernährungs-, und Gesundheitsforschung in Europa harmonisiert zusammenzuführen. (siehe auch: <http://edepot.wur.nl/444604>).

### **Einsatz der Bundesmittel für die Erreichung des geplanten Vorhabenziels**

Der Einsatz der Bundesmittel war entscheidend für die Erreichung des geplanten Vorhabenziels. Die BMEL Fördermittel für das Projekt konnten genutzt werden, um Daten aus an der TU München durchgeführte Humanstudien (die alleine einen Wert von ca. 600.000 Euro repräsentieren) zusammenzustellen, zu formatieren, adäquat zu dokumentieren und in die Datenbank von DASH-IN einzubringen. Dies setzte auch einen beständigen Austausch mit den Partnern in elektronischer bzw. telefonischer Form und in speziellen Arbeitstreffen des ENPADASI-Konsortiums voraus. Die Initiative ermöglicht auch der TU München den Zugang zu einem wichtigen Paket von Befunden aus Interventionsstudien.

### **Präsentationsmöglichkeiten für mögliche Nutzer**

Die Datenbank sowie die Analysesoftware von DASH-IN aus ENPADASI stehen den Projektpartnern zur Verfügung und erlauben eine Einbindung eigener Datensätze in die in der Datenbank hinterlegte Grundstruktur. Dabei bietet die Analysesoftware verschiedenen Analysemöglichkeiten, die einen Vergleich sowie eine Integration mit vorhandenen Datensätzen möglich machen. Die Hintergrunddaten sowie die Ergebnisse dieser Analysen stehen den ENPADASI Partnern zur Verfügung und können für Präsentationszwecke genutzt werden.