

Data use for Access & Participation in higher education

Review and recommendations by
CFE Research for the Office for Students

© Office for Students 2020



Report author(s):

Joanne Moore, CFE Associate
Dr Sarah Tazzyman
Lindsey Bowes
Dr Guy Birkin

For more information about this report please contact:
Lindsey Bowes

lindsey.bowes@cfe.org.uk

CFE Research
Phoenix Yard
Upper Brown Street
Leicester
LE1 5TE

0116 229 3300

www.cfe.org.uk

© Office for Students 2020

Established in 1997, CFE is an independent not-for-profit company specialising research and evaluation services in education, wellbeing and the economy.

Contents

Glossary	2
Executive summary	7
01. Introduction	10
Background and context	10
Research aims and objectives	11
Structure of the report	14
02. Overview of the current data landscape	15
Introduction	15
Data sources	15
Datasets for targeting.....	21
Datasets for Monitoring.....	33
Datasets for evaluation.....	40
03. Engagement with tracking organisations and value for money	55
Introduction	55
Operating context and services provided by trackers	55
Use of tracking services to support targeting, monitoring and evaluation.....	60
The cost of tracking.....	64
Discussion and future models.....	70
04. Practice-related issues	77
Introduction	77
Barriers and solutions.....	77
05. Legislation and ethical considerations	84
The General Data Protection Regulations	84
06. Alignment of schools and higher education	91
Introduction	91
Indicators and measures.....	91
Opportunities for shared monitoring and reporting	94
Discussion and future models.....	95
07. Summary and recommendations	97
Higher education providers' engagement with data.....	98
Appendix 1: Methodological note	105
Appendix 2: Examples from other sectors	108
Appendix 3: Overview of the functionality of the tracking services	112
Appendix 4: Barriers to effective targeting, monitoring and evaluation	119

Glossary

Key Terms

Terms in bold text within the definitions below have their own entries in the table.

Administrative data	Information routinely gathered as part of organisations' day-to-day activities. In the context of access and participation, this includes schools' data on pupil characteristics, attendance and attainment, such as that captured in the National Pupil Database (NPD)
Aimhigher	A national programme funded that ran from 2004 to 2010 with the aim of widening access in higher education for pupils from less advantaged backgrounds
Data controller	The individual or organisation responsible for collecting data and/or determining how it is processed
Data processor	The individual or organisation responsible for processing data in accordance with instructions from the data controller
Experimental method	A research design that can determine causality, based on random assignment of participants to test and control groups
Intervention	An activity or programme delivered as part of outreach
Monitoring	Regular, systematic observation of progress against targets
Outreach	Activity by higher education providers to encourage participation of prospective students, especially those from under-represented groups
Personal data	Information that enables the identification of an individual person
Quasi-experimental methods	A research design that estimates causality, like an experimental method , but where random assignment of participants to groups is not possible, and alternative methods of assignment are used
Targeting	Identification of learners through the use of specified characteristics

Abbreviations

ACORN	Classification that segments the UK population: https://acorn.caci.co.uk/
AHE	Adult higher education rate
A&P	Access and participation
APP	Access and Participation Plan
AWM	Aimhigher West Midlands tracking service
BAME	Black, Asian and minority ethnic
CDS	Common Data Set initiative
CHE	College of Higher Education
CJSM	Criminal Justice System eMail secure email system
CRM	Customer relationship management
DARS	Data Access Request Service
DfE	Department for Education
DLHE	Destinations of Leavers from Higher Education survey
DWP	Department for Work and Pensions
EBS	Education Business System
EMWPREP	East Midlands Widening Participation Research and Evaluation Partnership tracking service
ESFA	Education and Skills Funding Agency
EXACT	EXACT is a UCAS Media data service that can deliver datasets to a particular specification: https://www.ucas.com/data-and-analysis/data-products-and-services/exact
FSM	Free school meals
FTE	Full-time equivalent [employment]
GDPR	General Data Protection Regulations

GSI	Government Secure Intranet
HEAT	Higher Education Access Tracker tracking service
HEFCE	Higher Education Funding Council
HEI	Higher Education Institution
HEIDI Plus	Higher education business intelligence by HESA
HES	Hospital Episode Statistics
HESA	The Higher Education Statistics Agency
HMRC	HM Revenues and Customs
ICO	Information Commissioner's Office
ID	Identification
IDACI	Income Deprivation Affecting Children Index: https://opendatacommunities.org/def/concept/general-concepts/imd/idaci
IMD	Index of Multiple Deprivation: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019
ILR	Individualised Learner Record: https://www.gov.uk/government/collections/individualised-learner-record-ilr
JDL	Justice Data Lab
JSAS	Justice Statistics Analytical Services
KPI	Key performance indicator
KS2 / KS3 etc.	Key Stage 3, etc.
LEAP	Learner Evaluation and Progression toolkit
LEO	Longitudinal Educational Outcomes data: https://www.gov.uk/government/statistics/graduate-outcomes-for-all-subjects-by-university
LPN	Low participation neighbourhoods

MEM	Multiple Equality Measure
MHCLG	Ministry of Housing, Communities & Local Government
MIS	Management information system
MoJ	Ministry of Justice
MSOA	Middle layer super output area
NCOP	The National Collaborative Outreach Programme (Known as Uni Connect from February 2020)
NEON	the National Education Opportunities Network
NPD	National Pupil Database: https://www.gov.uk/government/collections/national-pupil-database
NSS	National Student Survey
OEP	Online Events Programme
OFFA	Office for Fair Access
OfS	The Office for Students
ONS	The Office for National Statistics
PDS	Personal Demographics Service
PG	Postgraduate
PNC	Police National Computer
PNN	Police National Network
POLAR	Participation of Local Areas, a geographical measure of higher education participation: https://www.officeforstudents.org.uk/data-and-analysis/young-participation-by-area/
PROMs	Patient Reported Outcome Measures
PSM	Propensity score matching
PVC	Pro-vice-chancellor
SAP	Student Ambassador Portal

SEN	Special educational needs
SITS	Strategic Information Technology Systems, a student records management system
SLC	Student Loans Company
STROBE	UCAS service that can track individuals into the UCAS applications system, and report anonymously on their outcomes or characteristics at aggregate levels: https://www.ucas.com/data-and-analysis/data-products-and-services/strobe
TASO	Transforming Access and Student Outcomes in higher education, a new 'What Works' centre for higher education.
TEF	Teaching Excellence and Student Outcomes Framework. For more information go to: https://www.officeforstudents.org.uk/advice-and-guidance/teaching/what-is-the-tef/
Tribal EBS	EBS is a student information system by Tribal
UCAS	Universities and Colleges Admissions Service
UPN	Unique Pupil Number
WP	Widening participation

Executive summary

Introduction

Data plays a crucial role in the effective planning, delivery and evaluation of access and participation (A&P). It can support higher-education providers to identify priority groups, monitor progress and engagement, and understand the effectiveness of different types of interventions at each stage in the student lifecycle. Higher-education providers currently use a variety of data types and systems to inform the targeting, monitoring and evaluation of A&P activities. However, little is known about what data is used at the different stages, how it is used and how useful it is. This report goes some way to addressing these gaps and identifies the barriers and enablers of effective data use in this context. The infrastructure and functionality required to improve access to, and use of, data is also explored.

The report's findings and recommendations are informed by desk research and consultation with data providers, tracking services and data users in the higher-education sector. The insights are designed to help the Office for Students (OfS) to develop guidance and resources to support effective data use as it works with the sector to achieve the goal of eliminating gaps in participation between the most- and least-represented groups within 20 years.

The higher-education data landscape

Providers draw on a range of data types, systems, indicators and measures to support the targeting, monitoring and evaluation of their A&P activities. The stage in the student lifecycle, wider institutional factors and external drivers, such as cost and the timeliness and robustness of data, all impact on what data is accessed and how it is applied.

Many higher-education providers use the services offered by three tracking organisations to overcome some of the challenges of targeting, monitoring and evaluating outreach activities. These organisations support members to track participation in outreach through data collection and the provision of an online database. Tools and resources to enable higher-education providers to effectively monitor and evaluate outreach are also provided.

Key findings

The data landscape is complex and this presents a barrier to effective data use. The array of datasets and measures is challenging to navigate and makes it difficult for higher-education providers to identify which are the most relevant for a given purpose. In a diverse sector, there are multiple data 'owners' with different conditions of access and use which can inhibit access to optimised data. The different systems adopted by data providers can also impact on the ability of higher-education providers to obtain and analyse data.

Data use is often related to the size and strategic direction of higher-education providers. Resourcing issues, including lack of staff time, and structural constraints, can often result in sub-optimal use of data in the targeting, monitoring and evaluation of A&P work. Skills gaps also limit some higher-education providers' capacity to obtain, analyse and interpret data.

Barriers in targeting, monitoring and evaluation

Issues of consistency, availability and quality of data for targeting, monitoring and evaluation are common at each stage in the student lifecycle.

Access to reliable and timely individual-level data for **targeting** is a major concern for higher-education providers. Data at this level is often incomplete and there is a risk that some individuals are not identified and miss out on support. Some higher-education providers examine intersections in multiple data sources to mitigate this. However, there is a risk of ‘double-counting’. Composite measures, such as the UCAS multiple equality measure (MEM) and HEAT Groups model, have been developed to address this.

A lack of templates and consistency in how A&P activities are defined and categorised act as barriers to accurate **monitoring** of A&P activities. The costs associated with using a tracking organisation is a barrier for smaller and specialist higher-education providers.

A lack of access to appropriate data presents a barrier to effective **evaluation** of A&P activities and results in inconsistent approaches across the sector. Assessing outcomes from A&P interventions against a comparison group also presents a challenge, compounded by a lack of understanding of experimental methodologies.

Legal and ethical considerations

The General Data Protection Regulations (GDPR) have had implications for the collection, use and storage of data. While the use of ‘consent’ as a legal basis for collecting and sharing data ensures the greatest control for data subjects, it can be more problematic than other legal bases for higher-education providers because of the administrative processes involved.

Aligning school and higher-education data

There are opportunities for closer alignment between school and higher-education data, with schools taking more responsibility for data that could be used for targeting and monitoring purposes. An integrated model of targeting, monitoring and reporting at the school level would, however, need support from the DfE. Tracking organisations could help to secure school engagement and promote shared monitoring and reporting.

Recommendations

Recommendations for the OfS to support improvements in existing data systems and infrastructure and the use of data by higher-education providers in the short- (6 months), medium- (18 months) and long-term (2–3 years) have been identified.

To support effective **targeting**, we recommend that the OfS:

- provide guidance about the use of different indicators and measures, and how this can support effective targeting [short term];
- provide case studies illustrating effective use of the A&P dataset [short term];
- provide options for a planning dataset for mature learners [short/medium term];
- provide a new resource for school-level targeting; and

- develop a methodology for using pupil-level datasets to prioritise learners for inclusion in outreach, using the OfS’s analytics¹ [short/medium term].

To support effective *monitoring*, we recommend that the OfS:

- clarify data requirements to monitor the effectiveness of outreach [short term];
- consider ways to minimise the cost to providers of accessing linked tracking data [short term];
- produce GDPR guidance on data collection for higher-education providers [short/medium term];
- work with stakeholders to develop consistent definitions of A&P activities, standard monitoring reports and dashboards [medium term];
- support access to linked data and establish the basis for the sharing of administrative data at an individual level [medium term];
- explore options for closer integration of schools and higher education datasets [medium/long term]; and
- explore, with DfE the development of a single student identifier [long term].

To support effective *evaluation*, we recommend that the OfS:

- produce guidance on how tracking data can support evaluation [short term];
- explore options for how linked National Pupil Database (NPD) data could be used in tracking systems, and to derive matched comparison groups [short/medium term];
- develop a toolkit to evaluate success activities [short to medium term];
- explore the feasibility of a pooled outreach activity dataset linked to matched data for tracking outcomes [medium/long term]; and
- consider options for centralising data linking and reporting participant outcomes, as part of an OfS function, or via a third party (e.g. ‘Data Lab’) [long term].

To support providers with *legal considerations*, we recommend that the OfS:

- work with the Information Commissioner’s Office (ICO) to clarify the legal basis for data collection, processing and matching, and outline when matched data can be shared with data providers to support evaluation activities [short term]; and
- develop best practice guidance for higher-education providers, informed by the DfE data protection toolkit [short/medium term].

To align school and higher education data, we recommend that the OfS:

- encourage higher-education providers to make A&P measures, such as look-up tables, more accessible to stakeholders and students [short term];
- develop a methodology to support schools to target groups at risk of not fulfilling their higher-education potential [short/medium term]; and
- promote wider use of the Unique Pupil Number (UPN) to encourage increased data-matching [long term].

¹ Composite measures are likely to be required that take account of attainment alongside factors of educational disadvantage. Further exploration of the experiences of providers using existing composite measures (e.g. UCAS MEM and HEAT groups) should inform this work, as well as best practice in ‘triangulating’ data.

01. Introduction

Here we set out the context and purpose of the review of the data landscape, including the research aims and objectives, and summarise the methodological approach.

Background and context

The Office for Students (OfS) was established in April 2018, replacing the Higher Education Funding Council (HEFCE) and the Office for Fair Access (OFFA) as the new regulator for higher education in England. It aims to ensure that “*every student, whatever their background, has a fulfilling experience of higher education that enriches their lives and careers*”.² A new outcomes-focused, risk-based approach to access and participation (A&P) will fulfil a central role in achieving this aim by challenging higher-education providers to eliminate the gaps in participation between the most- and least-represented groups across the entire student lifecycle within 20 years.

All higher-education providers that wish to charge fees above the basic level are required to set out their approach for improving equality of opportunity in higher education in an access and participation plan (APP). All APPs must detail how higher-education providers will seek to improve access and participation for five priority groups: (i) those living in area of low higher-education participation, or from lower-income/socio-economic status households; (ii) Black, Asian and Minority Ethnic (BAME) students; (iii) mature students; (iv) students with disability status; and (v) care leavers. Providers are also encouraged to consider how to close the gaps for carers, estranged students, people from gypsy, Roma and traveller communities, refugees, and children from military families. Higher-education providers are expected to obtain data on these under-represented groups, and draw on existing knowledge and evidence of effective practice, in order to develop an APP that sets out planned expenditure and associated targets and goals.

The OfS is also seeking a step change in the evaluation of A&P activities. The most recent regulatory notice³ and advice⁴ on preparing an APP both emphasise the importance of evaluation for driving continuous improvement and ensuring maximum benefit for target learners. Given the substantial investment in A&P by higher-education providers and

² OfS (2018) Office for Students Strategy 2018 to 2021.
<https://www.officeforstudents.org.uk/media/465d993d-daa8-42d2-a875-4a5fe63b211b/ofs-strategy-2018-21.pdf>

³ <https://www.officeforstudents.org.uk/media/obcce522-df4b-4517-a4fd-101c2468444a/regulatory-notice-1-access-and-participation-plan-guidance.pdf>

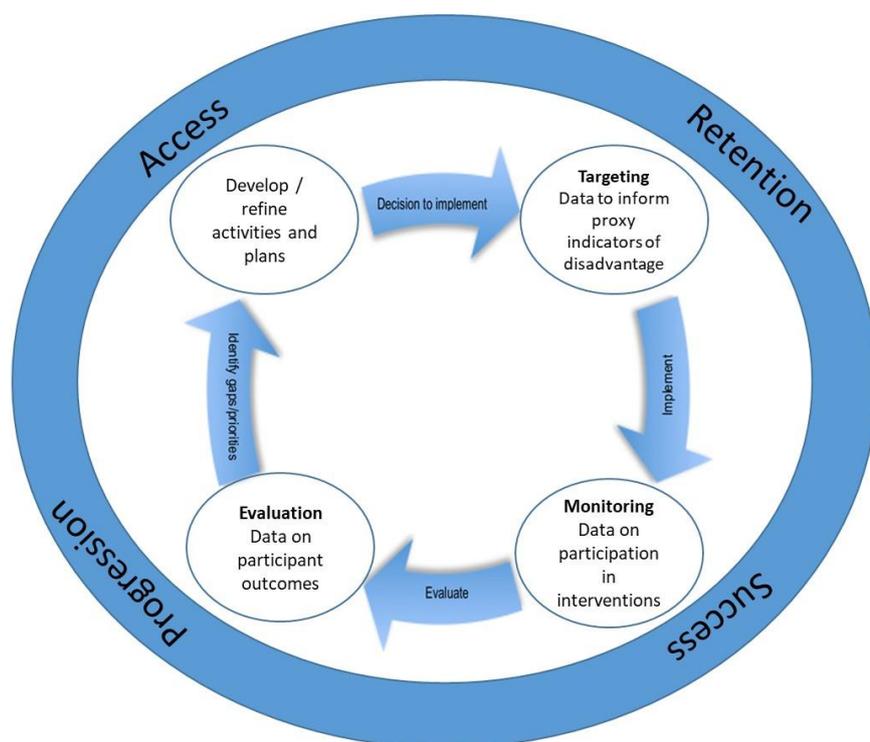
⁴ <https://www.officeforstudents.org.uk/media/1413599d-37bc-42ae-938a-d760d98c285b/regulatory-advice-6-how-to-prepare-your-access-and-participation-plan-guidance.pdf>

government each year,⁵ the increased focus on establishing ‘what works’ also reflects the OfS’s wider strategic objective to promote value for money for students and tax payers. Higher-education providers are, therefore, also expected to consider how the outcomes and impacts of their A&P work will be monitored and evaluated. The OfS has produced guidance and resources⁶ to help higher-education providers assess their current evaluation practice and support them to strengthen the evidence they produce.

Research aims and objectives

The OfS has identified great potential in the use of data to support the achievement of A&P objectives and for challenging individual providers on their performance. Data can fulfil a key role in: identifying priority groups/cohorts to ensure A&P activities are appropriately *targeted*; *monitoring* levels of engagement and the characteristics of participants to ensure activities are reaching the intended audience; understanding the effectiveness of different types of interventions at each stage in the student lifecycle; and capturing outcomes for learners. Crucially, the use of data can strengthen the evidence available to inform future planning and delivery, supporting a cycle of continuous reflection and improvement (see Figure 1).

Figure 1: Use of data in the targeting, monitoring and evaluation cycle.



⁵ According to the last monitoring report produced by OFFA, total investment in widening participation by higher-education providers was £883.5m in 2015/16

⁶ See <https://www.officeforstudents.org.uk/advice-and-guidance/promoting-equal-opportunities/evaluation-and-effective-practice/standards-of-evidence-and-evaluation-self-assessment-tool/>

A range of data is available to higher-education providers to support this process, but little is known about *which sources* of data are currently being used at each stage in the student lifecycle, *how useful* the available data is in this context and the factors that can act as *barriers to effective data use*. This research is designed to address these gaps in understanding and inform recommendations on how the OfS could work with other stakeholders to enhance the extent, nature and accessibility of data in future. The findings are also designed to inform the development of guidance for practitioners on the use of data for targeting, monitoring and evaluating A&P activities (to be published separately).

The aim of the research was to review and map the current data landscape and develop a fuller understanding of its functionality. In addressing these overarching aims, the research was designed to achieve the following objectives:

- provide an expert overview of arrangements for using data for targeting, monitoring and evaluating A&P activities;
- identify infrastructure and the functionality required of datasets to support effective practice;
- identify and assess opportunities, including alternative models, to develop the infrastructure and enhance data use in the future; and
- develop a value-for-money assessment framework.

Approach

This small-scale review focuses on the use of data as a mechanism for operationalising and understanding activities with priority groups or specific learners. It does not cover the use of data for research purposes more widely outside of the provider context, unless providers were seeking to test and contextualise sector-level research using their own datasets. In this context, data is defined as secondary statistical and other information sources including administrative data. This includes: data on pupil characteristics and educational attainment captured in the National Pupil Database; published area and school/college profiling data; data and information produced by schools and colleges and higher-education sector bodies (e.g. UCAS, HESA); tracking service data; and internal data collected by higher-education providers themselves (e.g. customer relationship management (CRM) and student information systems).

Methodology

The functional review was undertaken over six months using a mixed-methods approach that comprised three core elements: desk research, a stakeholder consultation, and data-user surveys.

Desk research

The desk research comprised three strands of activity:

- **A review of documentation**, including research, analysis, A&P consultation responses, sector guidance, funded programmes, strategies and plans of data providers and services, and best practice materials to provide insights into the policy

context, current arrangements for using data for targeting, monitoring and evaluation, and evidence of the strengths and weaknesses of current approaches;

- **A review of data protection legislation** and associated guidance and ethical guidelines to clarify the requirements for accessing and sharing data and explore the implications for current and future systems and processes; and
- **A review of alternative approaches** to using and sharing data in operation in other contexts, specifically health and criminal justice.

Stakeholder consultation

A total of 32 semi-structured telephone interviews were carried out with a range of senior stakeholders, including data providers, data users and representatives of the three tracking system providers. The interviews were designed to capture information about current arrangements for accessing and using data for the targeting, monitoring and evaluation of A&P. Levels of engagement, data flows and practical constraints, such as infrastructure, legal and organisational requirements, capability and organisational capacity were also explored, along with perceptions of the costs and benefits, and the value for money of current systems. Finally, the consultation considered future possibilities for the development of the data landscape to ensure effective use of data to promote equal opportunities in higher education.

Data-user surveys

The issues identified in the consultations with data users, along with insights from the desk research, informed the design of a data-user survey. The survey captured the views and experiences of staff involved in the use of data in targeting, monitoring and evaluating A&P activities in a range of contexts across the sector. Two online surveys were developed. The first focused on access to higher-education activities, the second on student success and progression.

Unique survey links were disseminated by the OfS on behalf of CFE to around 183 organisations listed on the OfS Register of English Higher-Education Providers database. The links were sent to a central contact who forwarded them to the appropriate members of staff within their organisation. A total of 82 unique higher-education providers responded to the survey - 76 responses to the 'access' survey and 60 to the success and progression survey - representing a 45 per cent response rate overall. The combined samples comprise a range of higher-education provider *types*, including 44 universities, 24 Further Education Colleges and 14 other types of provider. The samples also include providers of different sizes, defined by A&P 2019/20 projected expenditure.⁷

The analysis explored a range of data sources that higher-education providers use, the perceptions of their usefulness, and perceived barriers to data use for A&P targeting, monitoring and evaluation. Differences in perceptions and experiences were explored by provider size and tracking organisation membership status. Due to the small sample size, no

⁷ The sample was ranked in order of A&P projected spend (lowest to highest) and divided into three equal groups for the purposes of analysis.

significance testing was undertaken. Further details of the methodology are provided in [Appendix 1](#).

Structure of the report

The report is structured as follows:

- Chapter 2 outlines the types of data sources available and evidence on how they are currently accessed and used by higher-education providers.
- Chapter 3 assesses current arrangements for data use via tracking organisations, and discusses a value-for-money framework.
- Chapter 4 considers practice issues underpinning the use of data, in terms of barriers and constraints to effective practice.
- Chapter 5 discusses the legal considerations.
- Chapter 6 explores opportunities emerging from the review in terms of alignment of higher education and school data.
- Chapter 7 summarises the main findings and presents a series of emerging recommendations for the OfS.

Each chapter concludes with a discussion of future or alternative models.

Four appendices are included:

- [Appendix 1](#): Methodological note.
- [Appendix 2](#): Examples from other sectors.
- [Appendix 3](#): Overview of the functionality of the tracking services.
- [Appendix 4](#): Barriers to effective targeting, monitoring and evaluation.

02. Overview of the current data landscape

In this chapter, we map out the datasets and the mechanisms for accessing the different types of data that are currently used to target, monitor and evaluate access and participation activities in the English higher-education sector.

Introduction

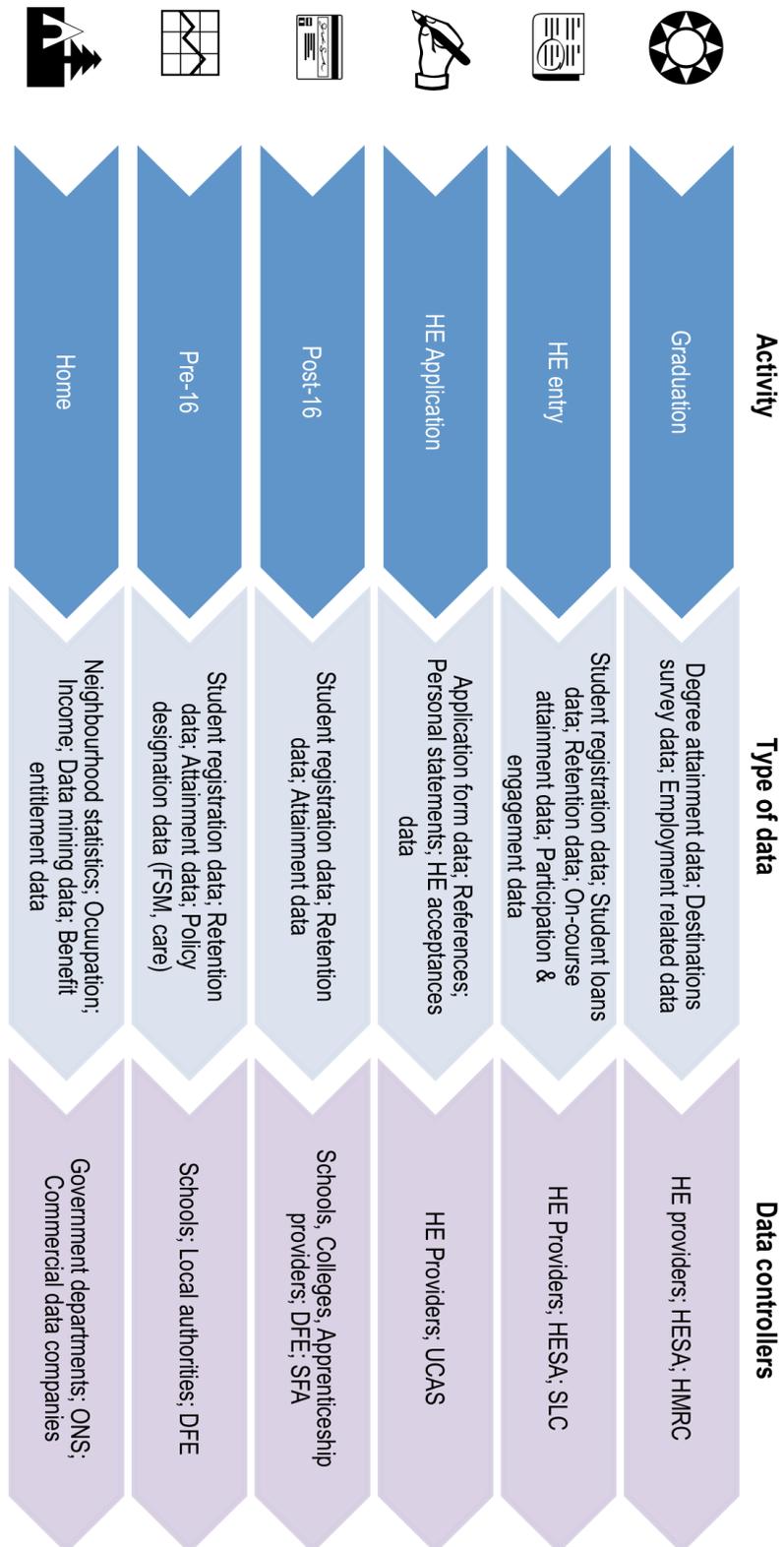
Mapping the data landscape for targeting, monitoring and evaluation of equal opportunities in higher education is challenging for a number of reasons. The landscape comprises a wide range of data sources, which include multiple indicators and measures. Data can be accessed and applied in different ways and at different levels depending on the purpose it is being used for, the stage in the student lifecycle and the strategic objectives of the higher-education provider. In a diverse sector, higher-education providers have different relationships with data providers and data intermediaries. They also use different systems and approaches for obtaining, analysing, interpreting and using data.

In this chapter, we first give an overview of the main sources of administrative data that higher-education providers can draw upon. We then draw on the primary qualitative research and survey findings to understand how this data is being applied at different stages and in different contexts and the issues that can arise.

Data sources

Administrative data sources can be distinguished according to whether the data is internal or external to higher-education providers. Within these categories, data can be public (i.e. published datasets) or private (i.e. access restricted to people in certain positions or by means of payment/membership). Sources can also be external (sourced through an outside body) or internal (available within the A&P provider organisation): for example, applicant data for a higher-education provider's own applicants (internal, private); published data on the population of applicants as a whole (external, public); and paid-for bespoke data (external, private). The main types of data available at each stage in the student lifecycle are illustrated in Figure 2 in relation to the stakeholders that control and process the data.

Figure 2: Main types of administrative data available



Data reveals background characteristics and progress through education to various providers

One of the reasons for data landscape complexity is that the same or similar data can be used for different purposes and accessed in different ways. Student data collected and held by higher-education providers forms the basis of data held and processed by data providers such as HESA, which is then shared back to providers in various forms, and is also used in different datasets, such as the OfS gaps analysis. In addition, through different approaches to matching and pooling data from these existing data sources, higher-education providers create new datasets to inform the targeting, monitoring, and/or evaluation of their A&P activities.

Linking across datasets is often desirable, particularly for monitoring and evaluation, e.g. bringing together data on participant contextual factors with activity-related data, or linking outcomes across educational phases. Practices of ‘fuzzy matching’ have been developed in order to link data and track learners across educational phases. The success of this approach relies on access to comprehensive personal data on which to base a match. Some data providers have developed services designed to facilitate more efficient data use by higher-education providers. For example, the UCAS contextual data service pulls together data on schools and matches it to individual applicants to enable contextualised admissions. The tracking services act as data intermediaries and source data from data controllers on behalf of members. Extensive secondary data fields including school and college level information is provided, together with postcode lookups and learner-based, composite targeting measures that can be matched to individuals. EMWPREP and AWM refresh local authority data to flag individual target groups within schools and colleges, including high-priority institutions to identify participants who could benefit from outreach activity. These types of initiatives help to save time and effort locally, so long as higher-education providers have trust in the provenance of the data and the means to apply it.

Levels of data aggregation

The extent and nature of the data needed depends upon the target groups being considered and the stage in the student lifecycle. For some groups/stages, data is required at the level of the individual (i.e. personal data). More details about the range of specific data sources and indicators that providers might usefully draw upon to inform their targeting of A&P activity can be located in the provider guidance document.⁸ In general, the level and type of data required for these different purposes can be mapped out as detailed below.

Area- and school-level data

Area- and school-level data is useful to inform high-level *targeting*, particularly at outreach and other pre-application stages, to contextualise applicants. Sources include published datasets at area and school/college level. As well as linking to measures of disadvantage (e.g. IMD), several datasets have been generated based upon historical patterns and trends in higher education: for example, UCAS progression data for schools and colleges, destinations data on the school performance tables, and POLAR maps. Area profiling data, together with school and college profiling data, is also available via the three tracking

⁸ <https://www.officeforstudents.org.uk/publications/regulatory-notice-1-access-and-participation-plan-guidance/>

systems (e.g. postcode lookups). This service can support planning, learner targeting and analysis of intervention-targeting effectiveness from administrative datasets.

Cohort-level data

Cohort data is useful for *monitoring* purposes but it needs to be matched to A&P activity/participation data in order to become an appropriate output measure of how groups are being supported. Higher-education providers looking to assess the effectiveness of their targeting at the cohort level may also want to monitor individual level factors/participation, and in a multi-activity programme, capture individual-activity intersectional data. This enables higher-education providers to explore the combination of activities that different groups of target students engage in, and can contribute to an understanding of the relative effectiveness of different types and/or intensities of engagement at the evaluation stage.

Individual-level data

A deeper layer of individual-level data and strategies such as contextual admissions help both outreach practitioners and student success and engagement teams to *target* specific sub-groups, such as care leavers or estranged students. To ensure the intended groups for activities are targeted, stakeholders can draw upon data on household socio-demographic factors (e.g. household income, family history of higher education, occupational classification); data applied to individuals to determine their eligibility for certain types of support (e.g. Free School Meals (FSM)) and/or data on personal characteristics (e.g. ethnic background). This data can be obtained from sources such as school/college records, the National Pupil Database (NPD), UCAS application forms, student registration data, outreach activity participation records, or by applying proxy indicators (e.g. matched to postcode). Individual-level data might also come into play at the *monitoring* stage, (depending on data quality issues discussed below) to ensure the targeting of activities has been successful. Table 1 (below) elaborates on the idea that datasets can be distinguished by who and what they measure, and shows the range of sources upon which providers draw to inform the targeting of their A&P work, focusing on external datasets rather than on internal recording systems. New datasets and mechanisms for using data are being made available. For example, the OfS A&P dataset is a recent addition to the data landscape, using HESA student returns data. We understand that the DfE is planning to make new data from the National Pupil Database (NPD) available via the school performance information. This development is designed to give stakeholders easier access to a new measure for schools based on the levels of higher-education progression.⁹

Factors influencing choice of data source

The approach used for targeting, monitoring and evaluation of both access and success and progression activities will be nuanced and will be adapted on the ground, not least because the various decisions made at different points in the student lifecycle will affect what and

⁹ Current Key Stage 4 measures include pupil destinations (education and employment after key stage 4), which can be compared with pupils' destinations with those of pupils at state-funded schools at local authority and national level.

how data is used. The following list of factors that appear to influence data usage is informed by our desk-based review and field work:

- The approach used to define and prioritise the target group(s) and set thresholds. It is important to note that different data providers define their target groups differently even within the same source, so the methods/thresholds set are important;
- How, and at what stage, for targeting, monitoring or evaluation purposes, any specific selection criteria are applied;
- The systems used for data capture, storage and processing (and whether these work in silos or interface with other data);
- The targets or KPIs designed for reporting purposes;
- The nature of the interventions (e.g. whether targeted and sustained activities are treated differently from untargeted or one-off interventions);
- The outcome and impact measures selected; and
- The evaluation methodology being used and the analytical strategy (e.g. whether aiming for data on comparators as well as participants).

Table 1: Summary of the higher-education data landscape.

Targeting			
Data Providers	Source	Level	End User
OfS, MHCLG, ONS, DfE, UCAS, local authority	POLAR, IMD, IDACI, further education and skills participation and achievement, UCAS EXACT, ACORN	Area	<ul style="list-style-type: none"> - Outreach managers - planning committees - outreach practitioners - collaborative delivery partners, student success & engagement managers - A&P managers, A&P committees and steering groups - OfS
DfE, ESFA, schools/colleges, local authority, UCAS	Levels of disadvantage by FSM, pupil premium, area level markers, performance (rates of attainment), location	Area (school / college)	
Data from application forms & questionnaires, student record system, learner analytics, OfS, UCAS, Common Data Set (CDS) Initiative	POLAR, IMD, geo-demographic profiling, higher education qualification, first in family to go to higher education, household income, UCAS EXACT	Individual (household / family)	
	FSM, pupil premium, care leaver, disability, attainment, 'at risk', CDS, A&P dataset	Individual	

Monitoring			
Data Providers	Source	Level	End User
<i>Internal</i> (institutional CRM, student record systems, learner analytics) <i>External</i> (HESA student data, DfE)	Eligibility criteria for participation, relative levels of disadvantage, household factors, experience of educational disadvantage, protected characteristics, target/policy designated group	Individual	<ul style="list-style-type: none"> - Outreach practitioners - collaborative delivery partners - student support officers - personal tutors
	Participation of learner cohorts with poor record of higher-education participation; student groups identified as being at risk of not fulfilling higher-education potential, HESA student data, NPD linked data	Group (cohort level)	
Monitoring & Evaluation			
Data Providers	Source	Level	End User
DfE, tracking organisations, schools/colleges, institutional follow-up processes	NPD linked data (HEAT), predicted and actual attainment over time, progression across educational phases, school/college data	Group (intermediate outcomes)	<ul style="list-style-type: none"> - Delivery practitioners & partners - A&P managers, committees & steering groups, decision makers - OfS
Evaluation			
Data Providers	Source	Level	End User
OfS, HESA, tracking organisations	A&P data, higher education applications and acceptances, Teaching Excellence and Student Outcomes Framework (TEF) metrics, sustained higher education entrants, retention and success in higher education, on-programme success measures, progression to graduate level jobs and further study, National Student Survey (NSS), Destinations of Leavers from Higher Education (DLHE), tracking data	Individual (outcomes level)	<ul style="list-style-type: none"> - OfS - institutional policy makers - DfE - Transforming Access and Student Outcomes in Higher Education (TASO) - Students - Advisers - External audiences

Institutional student records systems, learner analytics. UCAS, schools/colleges, HESA, tracking organisations	Retention and success in higher education, on-programme success measures, progression to graduate level jobs and further study	Group (gaps analysis)	
	Higher education applications and acceptances, sustained higher education entrants, retention and success in higher education, on-programme success measures, progression to graduate level jobs and further study, UCAS STROBE, school/college data, HESA student data, DLHE (HEAT)	Group (comparisons & controls)	

Datasets for targeting

The aim of effective targeting is to improve the participation of under-represented groups in higher education by ensuring that the ‘right’ groups and individuals benefit from A&P activity, but there is currently no single, ‘silver bullet’ indicator. There are limitations on targeting, not only because some datasets are difficult to obtain, but also because they are subject to interpretation or hard to apply to all applicants in a systematic way. It should be acknowledged that few datasets are 100% complete and accurate (even with ‘official’ data, some learners may not be included). Individual self-declared information from applicants may not be completely reliable since there is scope for applicants to misunderstand the question on the form or make false claims (requiring verification by providers).

The interviews with the institutional stakeholders confirmed the expectation that providers seek to draw on a range of data in order to direct their outreach activities towards specific target groups while aiming to be inclusive towards learners who fit their priorities. Only one higher-education provider survey respondent reported that they do not use data to target access activities, and one higher-education provider does not use data to target their success or progression activities. Insights from provider interviews suggest that, in general, an inclusive approach is taken to the delivery of the student success and progression interventions. Only a minority of these types of interventions appear to be targeted at specific types of students or groups.

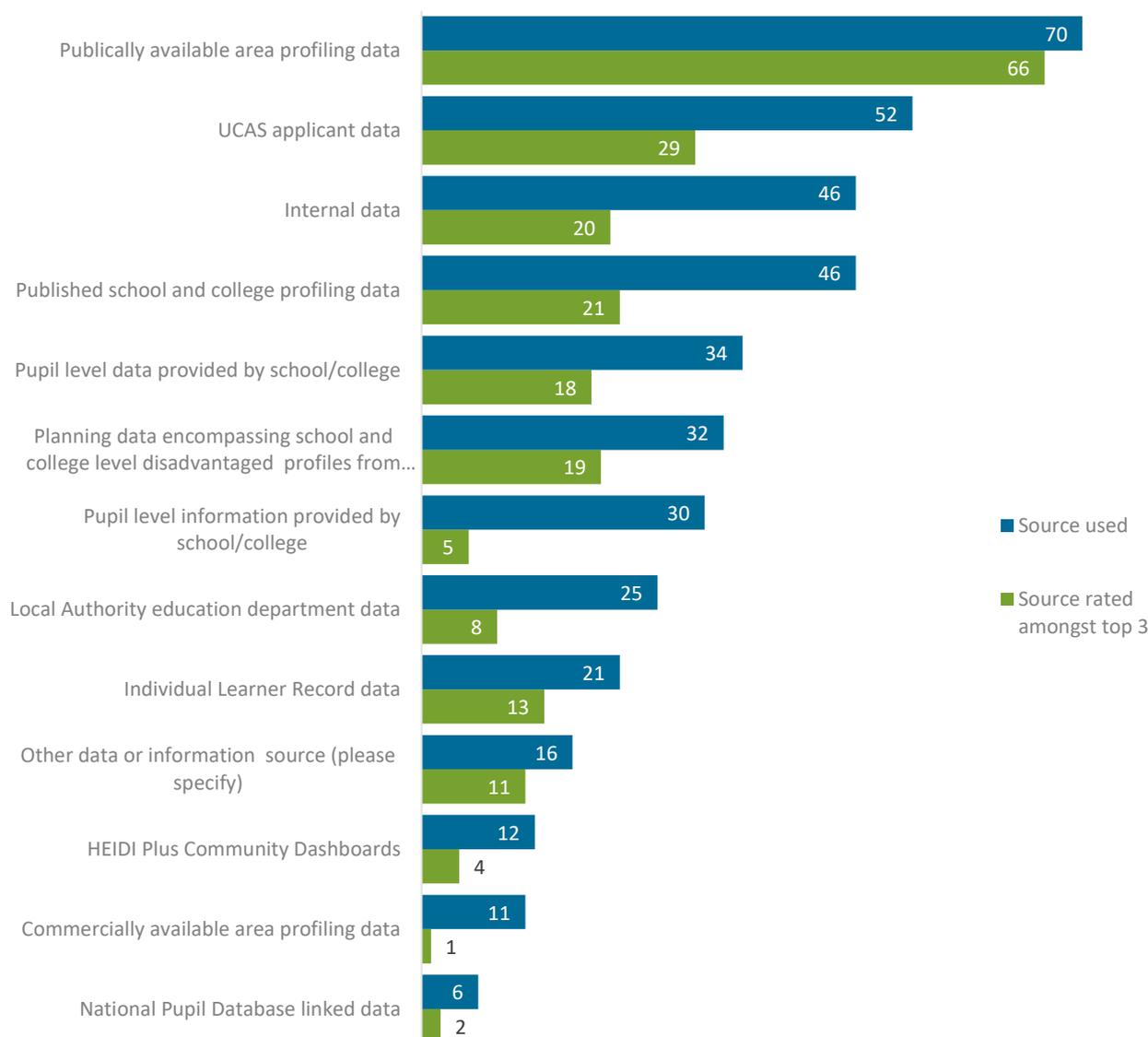
Access to data for targeting stands out as a major consideration for higher-education providers. Seven out of ten survey respondents said gaining access to appropriate datasets was a barrier to targeting access activities. Furthermore, access to data for targeting student success and progression activities was mentioned as a barrier by over a third of respondents. Even when a dataset is readily available, there may still be barriers in extracting useful data. In fact, almost two-fifths of respondents said that in relation to data

for targeting student success and progression activities, they had problems accessing the required data from within the dataset they were using.

External data sources for targeting A&P activities

The range of data sources and/or information that providers use to inform their A&P access-related activities is summarised in Figure 3. This clearly shows that the proxy measures required by providers to target their activities, and access activities in particular, are largely contained within external administrative datasets. It also shows that survey respondents perceive publicly available external datasets as the most useful to inform their targeting of A&P activities. The main types of external data sources for targeting A&P activities are reviewed in more detail below.

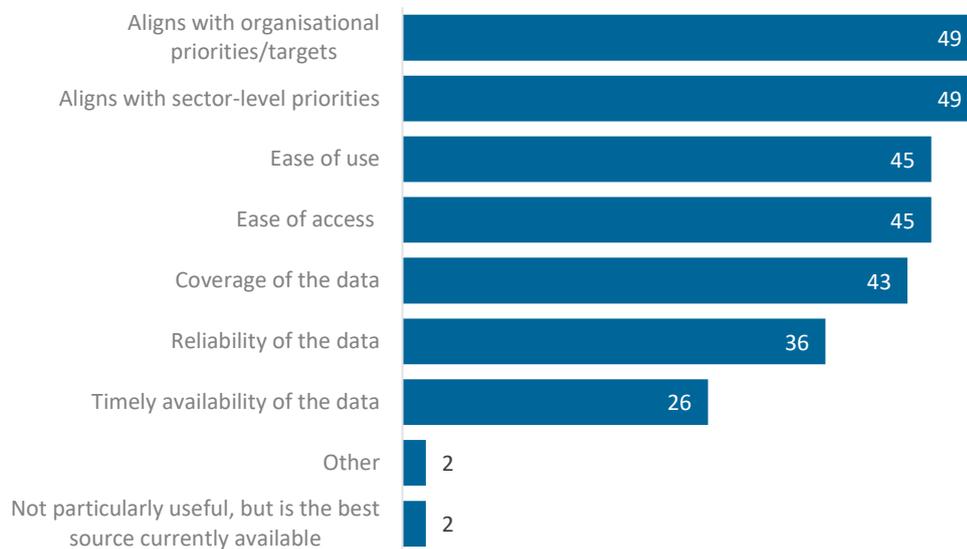
Figure 3: Use of sources of data for targeting A&P activities: Number of respondents who use each source and number who rank it in their top 3 sources (base = 76).



Area profiling data

When seeking to target access activities at the level of the cohort, publicly-available area profiling datasets (i.e. POLAR or IMD) are the most common source (used by nine out of ten survey respondents). This source is regarded as the most useful by the majority of providers (94%) for a range of reasons including ease of access, ease of use, good fit between the data and organisational and sector priorities, and good data coverage (Figure 4). Only a minority of survey respondents (n=11) said they use commercially-available area profiling data (e.g. ACORN). Further analysis reveals that larger providers (as defined by level of A&P expenditure) are more likely to make use of these commercial sources. None of the respondents with the smallest programmes by level of expenditure use commercial sources.

Figure 4: Reasons why publicly-available profiling data sources are useful (base = 66).



Area-based datasets (e.g. POLAR, IMD, ONS) can provide several benefits to enable effective targeting of A&P activity, including the relative ease of access in most instances (e.g. via postcode lookups) and the fact that much of the data is publicly available. One of the drawbacks is that the data may not be indicative of specific individual circumstances and this can increase the risk of false flagging, or missing target groups in highly-diverse population areas such as London. The desk research and qualitative research with stakeholders highlighted the well-rehearsed issues associated with POLAR measures. Strategies for addressing some of the weaknesses include applying data at lower levels of geographical concentration (as in the latest iteration of POLAR) and/or drawing on multiple data sources. Triangulating multiple sources of data can also help to ensure a more nuanced definition of ‘disadvantage’ and approach to targeting. It also helps to increase the validity and reliability of the measures used.

“Looking at LPN data, it’s not a robust measure of disadvantage. To use that as the basis of any kind of decision-making process means that there are students falling through gaps. There has got to be other measures in there, otherwise we’re

spending all this time and effort and there are students that we're missing just because they don't meet a specific dataset." [INT11]

School and college profiling data

Published school and college profiling data (e.g. performance tables) are used by six out of ten higher-education provider survey respondents to target their access activities. Higher-education providers with the largest A&P expenditure are most likely to use this source, presumably to factor attainment into their targeting decisions. Those who reported this as one of their most useful sources (n=21) explained that it was useful because it fitted with their institutional priorities. Good data coverage, reliability of the data and ease of access are the other main reasons highlighted by survey respondents (Figure 5). One third of survey respondents (n=25) use local authority education department data to target their access activities.

Figure 5: Reasons why published school and college profiling data is useful (base = 21).



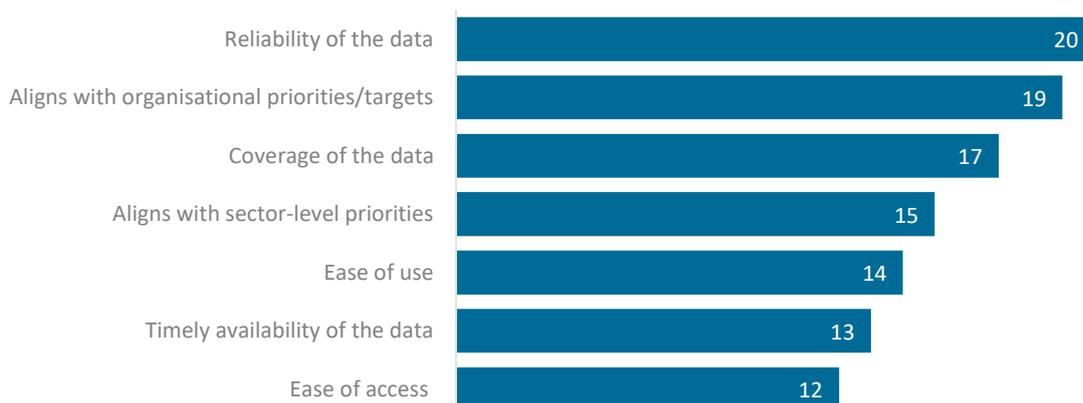
National Pupil Database (NPD) linked data is used by a minority of respondents (n=6) and most of these are higher-education providers with the largest A&P programmes by level of expenditure. Higher-education providers with the smallest programmes by level of A&P expenditure are least likely to use published school and college profiling data, and none use NPD. Most of these providers are colleges and specialist providers and, as a consequence, are more likely to use Individualised Learner Record (ILR) data (n=21).

Insights from the desk research and stakeholder interviews show that a key benefit of using school/college-level data is that it enables consideration of individual applicants in light of educational disadvantage/the circumstances in which their education attainment is achieved. Two of the main drawbacks identified through the current research are difficulties in interpreting school/college-level data (e.g. some learners may overachieve in underachieving schools) and inaccuracies from having to rely on historical data.

Higher-education sector data

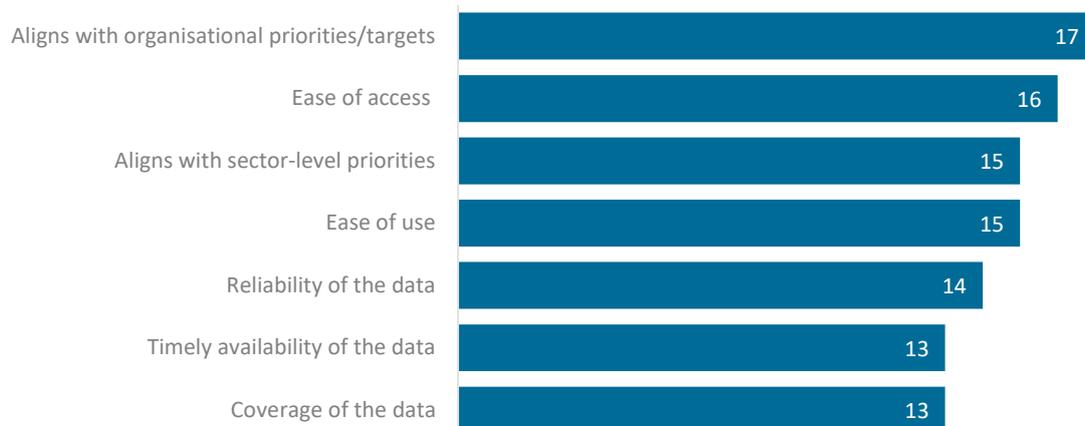
The survey results suggest that providers find it easier to draw on higher-education sector data than administrative data produced by the compulsory education system. Over two-thirds of survey respondents overall (n=52) use UCAS applicant data to support the targeting of their access activities, and this is rated as one of the top three most useful sources for over half (n=29) of those who use it. Reliability and fit with organisational priorities are the most prevalent reasons for rating this source as useful, along with data coverage (Figure 6). The timeliness of UCAS data is perceived as useful, although presumably the usefulness of the applicant data is in mapping past trends in progression rather than identifying current learners or access participation. The planned inclusion of a progression measure within the DfE performance tables may go some way to making progression data more widely available as part of the published suite of data on schools and colleges for targeting purposes. One of the main limitations of UCAS data is that it does not provide a measure of sustained higher education destinations as it is based on learners who have applied and decisions made during the application process. There is also a risk of missing applicants to higher education who apply through non-UCAS routes.

Figure 6: Reasons for use of UCAS data (base = 29).



Tracking services support targeting of A&P activities in a range of ways, including through the provision of access to new data, lists, and look-ups. In their role as data intermediaries, they also facilitate access to contextual data in the NPD. A total of 41 respondents to the survey work for higher-education providers that are members of a tracking service. Almost three-quarters of members (n=32) use the tracking service to source planning data encompassing school and college-level disadvantage profiles to inform targeting. Respondents reported that this planning data was useful because it helped them meet their organisational (n=17) and sector-level priorities (n=15) (Figure 7). The postcode look-up feature provided by the tracking services also enables providers to target students for specific events via their postcode data. The ease of access, ease of use and reliability of this data are also perceived to be useful features. See Chapter 3 for a fuller discussion of the role of the tracking services.

Figure 7: Reasons for use of planning data (encompassing school- and college-level disadvantaged profiles) from a tracking organisation (base = 19).



A minority of survey respondents (n=12) said they used the HEIDI Plus community dashboards to target their access activities. Alignment with organisational and institutional priorities are perceived as useful features of this data source, although only a minority said they were easy to access and use.

Other higher-education sector data is commonly used in the targeting of student success and progression activities in particular. For example, the majority of respondents use Teaching Excellence and Student Outcomes Framework (TEF) metrics (n=42, 70%) and National Student Survey (NSS) (n = 39, 65%) data for this purpose, although only a minority report that these were the main sources they used. Alignment with sector and institutional priorities, and ease of access and interpretation, mean TEF metrics are regarded as particularly useful for targeting. The NSS is perceived to be useful because it provides detailed, consistent data based on high response rates.

Nine out ten survey respondents (n= 54) said they used the POLAR¹⁰ to target student success and progression activities, and this was a secondary source in all cases. The qualitative research with institutions suggests that the importance of this dataset will probably increase in future. Practitioners particularly welcomed the opportunity to use the same data as the regulator and gain access to new targeting indicators that could provide the opportunity to use data across the student lifecycle. However, there are some concerns about the indicators of disadvantage used and the impact that this might have. The main barriers mentioned are practical constraints such as having the right expertise in place to work with complex datasets and a lack of time and resources.

School/college-provided data

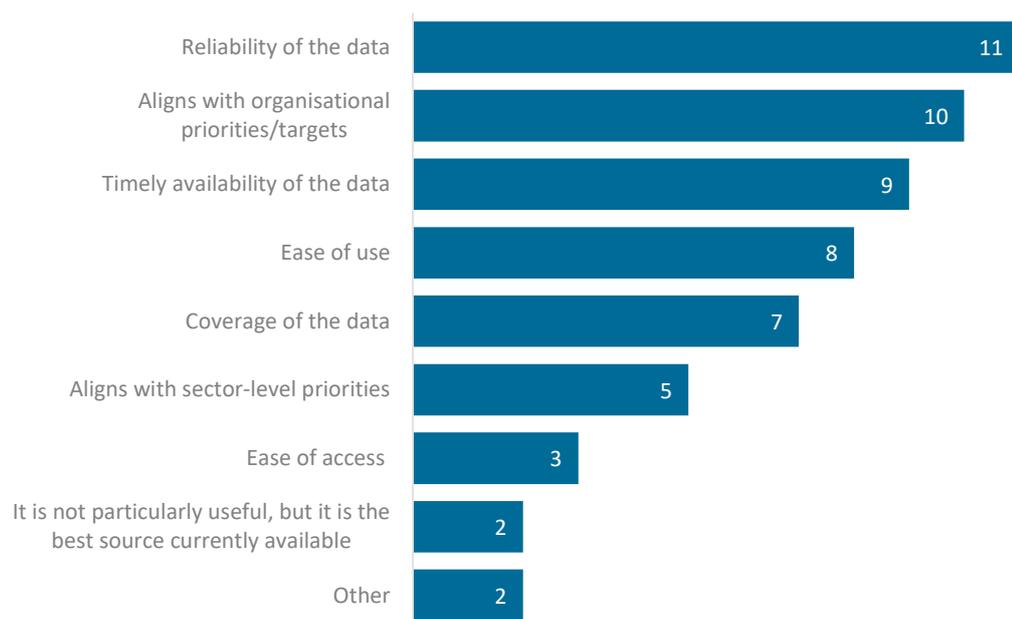
The use of individual data is frequently preferred by some because it is specific to the person concerned, is often validated and highly reliable. However, gaining access to individual data

¹⁰ The dataset is available here: <https://www.officeforstudents.org.uk/data-and-analysis/access-and-participation-data-dashboard/>

from administrative datasets, together with the timeliness of the data, are common barriers. When it comes to *individual-level* targeting, interviewees spoke about their experience of addressing the limitations of administrative data (e.g. NPD, ILR) by drawing on data and information from partner schools and colleges involved in their access to higher-education work. The tracking services are also supporting higher-education providers with individual-level targeting. For example, EMWPREP has sourced data for higher-education providers and National Collaborative Outreach Programme (NCOP)¹¹ partnerships from several local authorities in order to prepare priority lists of pupils that schools can use to target individuals for inclusion in local outreach activities. AWM provide a similar service by sourcing data from schools and colleges on an annual basis to support higher-education providers to target individuals for specific interventions (see Chapter 3 for more details).

More than two-fifths of survey respondents (n=34) draw on pupil-level information provided by schools/colleges to support individual-level targeting of outreach. This is one of the most useful sources of data for just over half (n=18) of those who use it. 40% (n=30) use pupil-level information provided by teachers, but this is perceived as less useful. Pupil-level data is regarded as useful because it is considered to be reliable and timely and because it fits with their organisational priorities (Figure 8).

Figure 8: Reasons why pupil-level data is perceived as being useful (base = 18).



Higher-education providers with the largest programmes by level of A&P expenditure were more likely than those with smaller programmes to draw on pupil-level *data* provided by schools and colleges. However, higher-education providers irrespective of A&P expenditure use pupil-level *information*, such as teacher feedback, provided by schools and colleges. Getting data from participants is the most prevalent barrier to effective use of data for

¹¹ The National Collaborative Outreach Programme was rebranded ‘Uni Connect’ in February 2020.

targeting of outreach activities, identified by almost three-quarters (73%, n=54) of respondents, irrespective of A&P programme size.

Making effective use of individual data for targeting

Some people contest the appropriateness of using individual data as a targeting approach, regarding it as a ‘deficit model’ that risks stigmatising those who are targeted and excluding those who could also potentially benefit from the support. However, there are a range of data sources and/or bundle measures that can help to mitigate against a potential ‘deficit model’ as detailed below.

Free School Meals (FSM) data is a proxy for household income which is commonly used to target individuals. However, those who are just above the threshold for FSM are also likely to be disadvantaged and if activity is targeted on this measure alone, these individuals miss out. Furthermore, eligibility for FSM for low-income households on the borderline often changes over time, and those who are eligible for FSM do not always take advantage of it (with discrepancies depending upon local authority policies). Therefore, FSM is most useful as a measure of school-level disadvantage, or as a proxy for low-income households, when used as one of a bundle of measures to establish the number of learners from low-income households.

Using a bundle of measures is one way to overcome the inherent limitations of a single data source. This was the recommendation of the Social Mobility Advisory Group to support the expansion of the data available to enable universities to assess their work on social mobility,¹² and was included in guidance for the use of data for targeting in contextualised admissions.¹³ This approach has been adopted by some data providers including the OfS, in the development of the A&P dataset, and UCAS and HEAT in the development of composite measures, including the multiple equality measure (MEM) and HEAT groups model. AWM have provided a composite model service since 2008, comprising a basket of measures sourced from NPD/ILR, IMD, FSM, 16–19 bursary, POLAR4 and size of WP cohorts. Using a composite model enables higher-education providers to band schools into low, medium and high priority groups. Some of the higher-education providers consulted (including those that already draw upon several datasets for targeting including ACORN, IMD, IDACI, FSM, school performance data and HESA data) are examining the use of these composite measures, recognising their potential to “...paint a much richer picture”.

However, using multiple indicators also presents challenges because data from the different sources is often available at varying levels. For example, FSM is an individual-level measure of disadvantage while IMD or IDACI and POLAR are geographically-based. A substantial amount of time is required to marry up different datasets and high-level technical skills are needed to manipulate and analyse the information. Furthermore, there is a risk of double-

¹² Universities UK (UUK), 2016. Working in partnership: enabling social mobility in higher education – the final report of the Social Mobility Advisory Group. London: UUK. <https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Pages/working-in-partnership-enabling-social-mobility-in-higher-education.aspx>

¹³ Supporting Professionalism in Admissions (SPA) (2016). Contextualised Admissions briefing – Triangulation of contextual data: Building a clearer picture of the individual applicant. Cheltenham: SPA. <https://www.ucas.com/file/233176/download?token=zPkYBBDF>

counting. When using multiple measures, HEFCE's work to assess the overlaps between IMD and POLAR is useful as it reassures the user that these measures capture different characteristics.¹⁴

There is also a lack of available measures of educational disadvantage for mature learners, despite mature learners being a policy priority for the OfS. One interviewee said they were targeting adults on the basis of deprivation indices, and also as part of a strategy of targeting the parents of disadvantaged youngsters. However, this was problematic due to data becoming out of date quickly. Partnership work with the local authority had enabled information on deprived areas to be triangulated with data on higher-education participation, but this was locally-specific data, and the approach could not be replicated regionally or nationally. The issue of data for mature learners can continue post application, as indicators of educational disadvantage in the compulsory education phase will no longer be available within the most recent published datasets for mature learners. The reliability of indicators also tends to reduce due to the fact that the measures may no longer capture factors relevant to the person's experience of relative disadvantage: for example, postcode of current address not reflecting background factors influential to higher-education prospects.

Whatever external data source is being used, trust and ownership of the data emerge as important considerations from the stakeholder interviews. Feedback from the higher-education providers suggests that a lot of time and effort is invested in re-creating results when data is drawn from national and local datasets, and matching the analysis of one data source to another in order to increase provider confidence in the data and its fitness for purpose in terms of targeting A&P activities. A&P staff members 'double check' national figures to see how they compare with their own data in order to comment and respond to the national data. This research reinforces the existing literature on administrative data use which points to a trade-off between the ease of accessibility of data for targeting and how useful it is. How higher-education providers seek to manage this trade-off, and the choices they ultimately make, will depend on local capacity and capability, as well as organisational priorities and the scale of their A&P programmes.

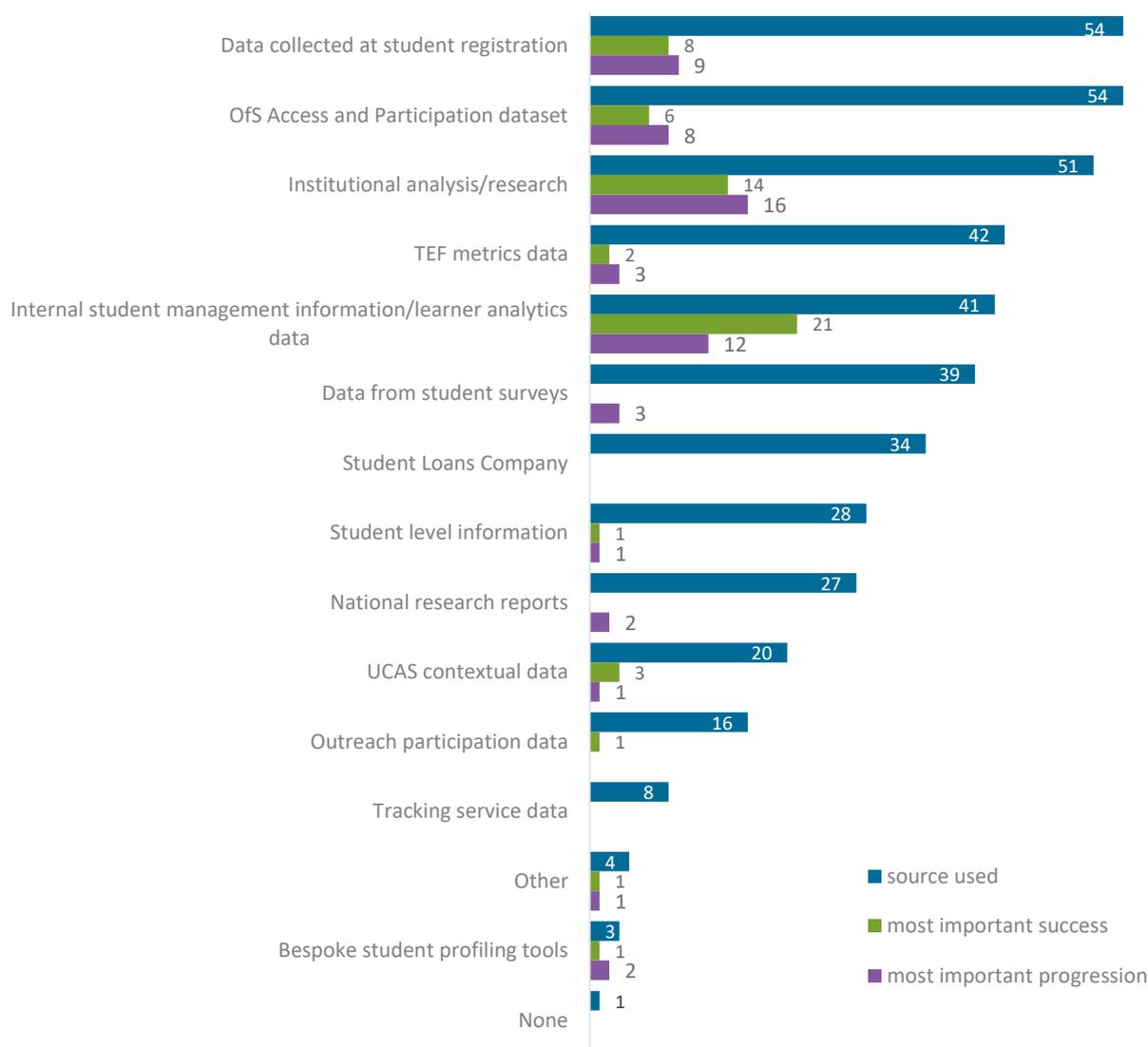
Internal data sources for targeting A&P activities

Internal data sources are predominantly used by higher-education providers to support the targeting of student success and progression activities. Student enrolment processes, application processes and sign-up forms are key sources of internal data that are designed to capture individual-level data for targeting purposes. The vast majority of survey respondents said they used data collected at student registration to target their success and progression activities, although this was rated as the most important source by only a minority (nine providers use it as the main source for their success activities and eight use it as the main source for targeting progression activities) (Figure 9). The benefits of sourcing data from the student registration process include the timeliness of the data, ease of access

¹⁴ HEFCE (2007). *Review of performance indicators: Outcomes and decisions*. Bristol: HEFCE. https://dera.ioe.ac.uk/6804/1/07_14.pdf

and data coverage. Given the fact that these processes are designed by the higher-education providers themselves, it is not surprising that this source is also regarded as useful because of its alignment with institutional priorities. Internal student management information systems (MIS)/learner analytics are used for targeting student success and progression activities by just over two-thirds of respondents (n=41) and it is perceived to be the most important source for supporting the targeting of A&P success activities to a greater extent than progression activities (Figure 9). As with student registration data, this source is perceived to be useful because of its timely availability, ease of access, coverage, and alignment with institutional priorities.

Figure 9: Sources of data used and most important sources for targeting Success & Progression activities (n=60).



The stakeholder interviews suggest, however, that the infrastructure in place within a minority of higher-education providers is not sufficient to provide timely data and that time

lags limit the usefulness of internal data for targeting in this context. One interviewee admitted that:

“...our own systems, most university systems, can’t produce up-to-date data.”
[INT16]

However, the extent to which these internal data sources are used to *directly* target individuals based on their characteristics (as with the targeting of outreach) for interventions is limited. The higher-education provider interviews reveal that it is more common for providers to *indirectly* target groups where wider indicators and trends suggest they are at risk of dropping out and/or failing to attain their full potential. Based on this data, cohorts of students on courses with historically high attrition rates or cohorts with certain types of prior qualification (e.g. vocational qualifications) could be targeted for support. Other risk factors that could lead to a student being targeted include persistent non-attendance, requests for extensions, missed deadlines or lack of engagement in facilities such as the library.

“We don’t target BAME students, we target the programmes they’re on. We look at programmes where we have a lot of students in that particular category and target the programme level, not the individual student.” [INT11]

“...We have a student dashboard which measures their electronic footprint - engagement with the uni - so we use the dashboard to target, not their characteristics... by targeting engagement we get to those who struggle.” [INT09]

Institutional analysis/research is an important source of data for targeting student success and progression activities, and 85% (n=51) of survey respondents use it. Of these, approximately a quarter identify it as their main source for success (n=14) and progression activities (n=16) (Figure 9). As with the other sources of internal data produced by the higher-education provider themselves, the majority of users report that what makes it particularly useful for targeting is the timely availability and coverage of this data, along with its ease of access.

Prior participation in outreach informs targeting of success and progression activities at 28% (n=16) of the higher-education providers who responded to the survey, but is the main source used for this purpose by just one provider. Prior outreach participation data can support a progression framework approach to widening participation activities that can help to ensure students receive the bespoke, tailored activities they are most likely to find beneficial. There are a number of limitations to using prior outreach participation data, which may explain why it is not used by more providers to target learners post-entry to higher education. Validating participation and completion in targeted programmes can be problematic and details of prior involvement in other programmes tends to be unavailable. There may also be a mismatch between the targeting criteria used by alternative schemes, and information about the targeting criteria used is frequently missing altogether.

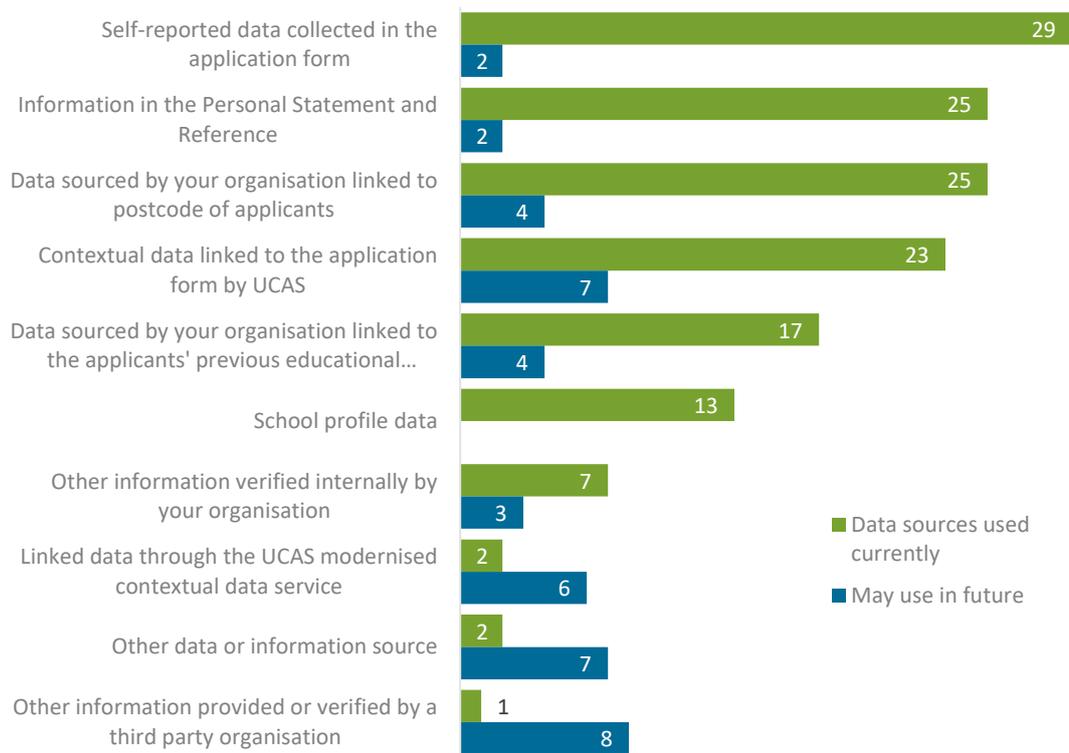
Feedback from the higher-education provider interviews suggests that key barriers to the effective use of data for targeting student success and progression activities include the

range of information available (often collected for different purposes) and differing views on the most appropriate data to use and how it should be interpreted for the purposes of informing A&P activities. Both the choice of data available and the methodologies for analysing and interpreting it are described as “overwhelming” by some A&P practitioners. In particular, as non-data analysts, they often find it difficult to discern the quality and robustness of the data available and whether it is appropriate to use it to target at the level of the department or course where sample sizes are small (see Chapter 4 for a more detailed overview of the barriers experienced by higher-education providers). In this context, what would help to enhance the use of data is further support along with the sharing of knowledge and insights on the most relevant data sources between access and student success and progression teams.

Contextual admissions

Contextual admissions are not an A&P intervention as such, but can be viewed as a strategy to support higher-education providers to achieve their A&P targets. Over half (n=42) of respondents to the survey said they used contextualised admissions as part of their wider A&P strategy and 39% (n=7) said they were planning to use data from the UCAS contextual data service in future. The complex interplay between data from a range of sources can be illustrated in the case of contextual admissions, as providers are seeking to build up a detailed picture by drawing on different sources. The most commonly-used sources to inform the contextual admissions process are self-reported data collected in the application form, information in the personal statement and reference, and data sources by providers linked to applicant postcode (Figure 10).

Figure 10: Which data sources are currently used to inform the contextual admissions process and which may be used in future (base =41).



Datasets for Monitoring

Data is commonly used by the majority of higher-education providers to monitor progress against A&P targets to understand how effectively interventions are engaging priority groups and/or meeting specific eligibility criteria. Data is also used for the purposes of monitoring the extent to which the needs of different priority groups are being met. Only four survey respondents (5%) do not routinely monitor their access activities. A similar proportion indicate that they do not monitor their student success or progression activities. The range of data sources that higher-education providers use to monitor participation in access, and success and progression, activities, and the extent to which they are perceived to be useful and important, are shown in the figures below.

Figure 11: Data systems used to monitor A&P access to higher education-related activities: which systems are used and which systems are rated in the top 3 (base =75).

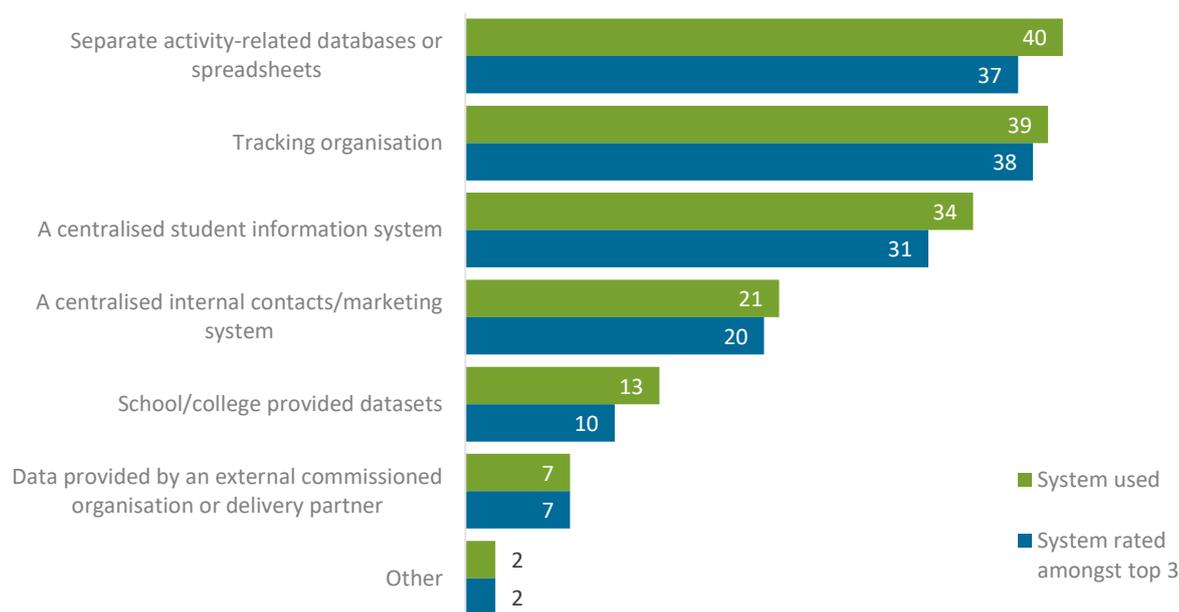


Figure 12: Data systems used to monitor A&P student success-related activities: which systems are used and which systems are most useful (base = 60).

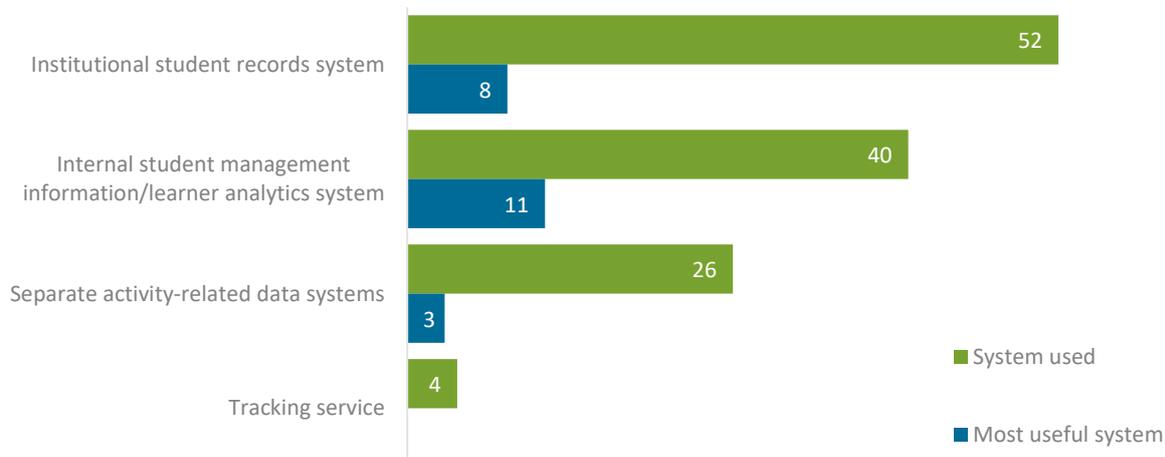
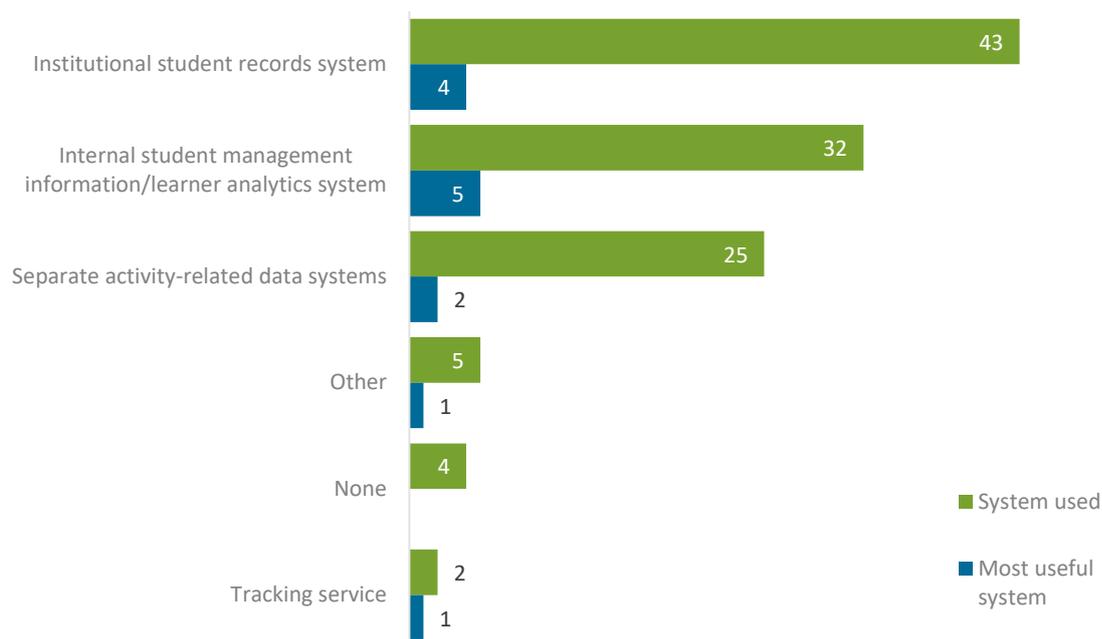


Figure 13: Data systems used to monitor A&P student progression-related activities: which systems are used and which systems are most useful (base = 60).



External data sources for monitoring of A&P activities

Monitoring is influenced by institutional values and strategies, as well as national policy and associated targets. Particular attention is paid to the measures that underpin the key performance indicators for A&P nationally, especially POLAR, and these are subsequently adopted by providers for their own monitoring purposes. The drawback of this approach is that some of the measures that inform national policy are not necessarily the most appropriate at the level of the higher-education provider; a more robust approach would involve selecting the indicator that is best suited to the task and the institutional context.

Among the members of tracking services who responded to the survey, just over half (n=39) use tracking data to monitor their access-related interventions. This is the main source of data used for this purpose by this group and is regarded as one of the top three useful data systems by almost all respondents (n=38). The majority of users of tracking services report that the ability to monitor participation across activities and identify participants in multiple interventions are the most useful features, along with the ease with which they are able to use the data to report against their targets and benchmark their performance. Two-thirds of respondents (n=25) also reported that tracking data was useful because it gave good coverage (See Figure 14).

Figure 14: Reasons why data systems are useful to inform the monitoring of A&P access-related activities (base = 38).



The tracking services are primarily designed to support access to higher education and provide few (if any) services specifically designed to support success and progression. However, a minority of respondents (n=4) indicated that they used tracking services as a source of secondary data for monitoring their success and progression activities (see Figure 12 above). The robustness of the data, data coverage, and ease of reporting were the reasons given as to why this source was perceived to be particularly useful.

External datasets are sometimes favoured at the monitoring stage because of the potential for benchmarking across the sector. Interviewees reported that they find both the HESA KPIs and the TEF metrics useful because they include comparable and consistent data that can be used for benchmarking purposes. A lack of knowledge and understanding about the data sources that are available for monitoring purposes was highlighted by one interviewee, and another described the problems they had encountered identifying suitable comparators for their provision (which involves flexible and part-time courses). The challenge of identifying suitable comparative data is compounded by the wide choice of methods and approaches for target setting, which reflect an individual provider's institutional priorities and might vary across different teams/departments, as well as the differing indicators available within datasets at different stages in the student lifecycle. Higher-education providers working with specific sub-groups such as mature and part-time learners also reflected on the paucity of data for monitoring and benchmarking performance in relation to A&P for those learners.

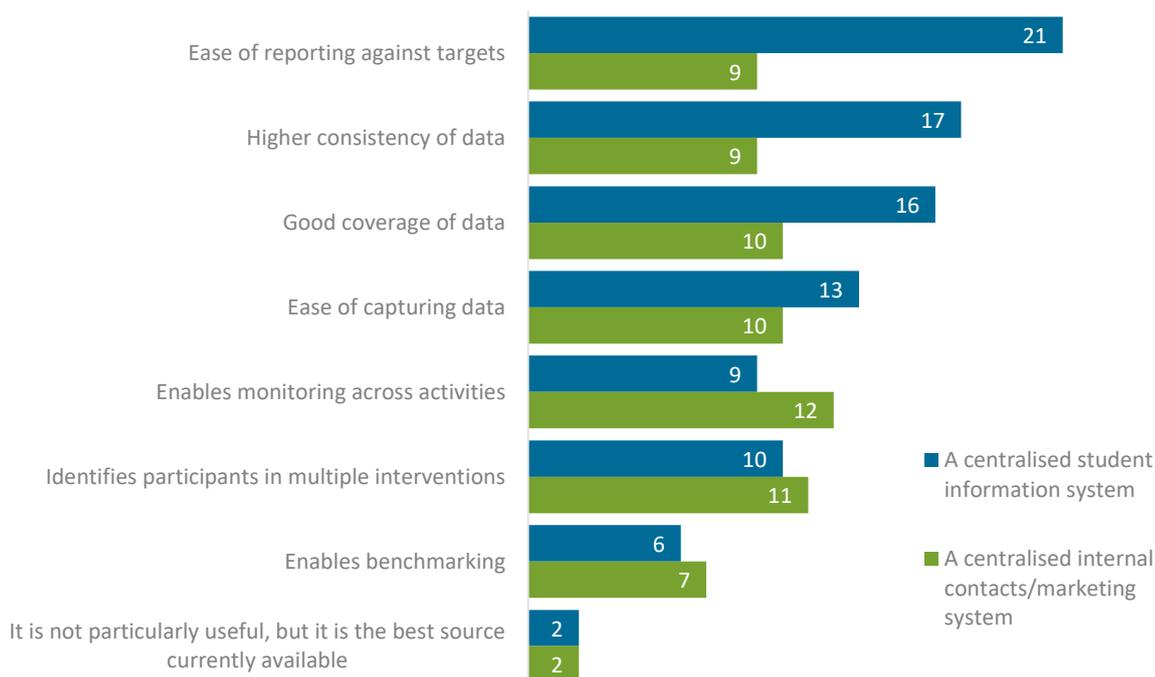
Issues of interpretation appear to cause problems for some practitioners, particularly when a large amount of data, including dis-aggregated data, is available. A key issue identified by interviewees was uncertainty about which metrics or gaps they should prioritise in the data. This issue is further compounded when different datasets provide different results for the same proxy measure. For example, one interviewee said that the provider took account of POLAR and IMD in its monitoring; while both are a measure of disadvantage, it does well

on one but not the other. This provider wanted clarity on which indicator to prioritise and was looking for guidance from the OfS.

Internal data sources for monitoring of A&P activities

Internal datasets developed and managed by higher-education providers are likely to include a wider range of measures than those contained in external administrative data sources because providers have the opportunity to collect additional data from learners, for example through application processes or student registration forms. Most providers use an online student registration system to streamline the process and inform subsequent monitoring. This can offer several advantages including centralised data storage, easy access and integration with other data such as fees collection, attendance and marking systems (Figure 15). Online registrations systems are designed to complement the requirements of other data providers such as HESA, which enables them to collect additional data relevant to their A&P strategic aims and objectives. Registration data can be particularly useful for monitoring the student journey since it enables nuanced analysis of patterns across student groups (in some instances linked to a learner analytics approach).

Figure 15: Reasons why ‘a centralised student information system’ and ‘centralised internal contacts/marketing system’ are useful to inform the monitoring of A&P access-related activities (base = 20–31).



The Career Registration project,¹⁵ funded as part of the HEFCE ‘measuring learning gain’ programme, is an example of cross-institutional work to identify dependable registration data collection methods that inform understanding of the support needs and career trajectories of different groups. A comprehensive dataset showing starting points and

¹⁵ More information at: <https://london.ac.uk/the-careers-group/research-unit/careers-registration>

progress was produced after piloting data collection, together with a series of recommendations for the implementation of Careers Registration in institutions and a standardised methodology/standardised information-capture on all students along a timeline.¹⁶

Once learners become a student at a particular provider, the provider is able to access more comprehensive, high quality data on the individual, such as data relating to household residual income. For example, the Student Loans Company (SLC) takes a rigorous approach to identifying household income, which has the potential to be used to target low-income background students (with caveats) for A&P activities, once on-course, and monitor their success and progression. The OfS is planning to use this data in the A&P dataset. Providers could also make more use of income data for research purposes to test the accuracy of other measures applied to learners at the pre-entry stage.

A common approach to monitoring *access* used by just over half of survey respondents (n=40) is the development of separate activity-related databases or spreadsheets. This is perceived to be one of their most useful data sources. However, almost a quarter (n=9) report that although it's not particularly useful for monitoring purposes, it is the best source currently available (Figure 16). The qualitative research suggests that some providers use a range of internal (bespoke departmental databases and institution-wide systems) and external systems, including one or more of the tracking organisations, consecutively.

Figure 16: Reasons why separate activity-related databases or spreadsheets are useful to inform the monitoring of A&P access-related activities (base = 37)



Over two-fifths of survey respondents use separate activity-related databases to monitor student success and progression activities. Many of those who used this as their main source of data for monitoring these activities did not regard this approach as particularly useful but was the only one currently available. The minority that favoured this approach

¹⁶ The data-collection process was also combined with a delivery aspect: as well as students picking one of 10 statements which they felt most closely reflected their current career thinking, respondents received a response with tips and ideas reflecting their year of study.

indicated that the main reason they found it useful was the robustness of the data. It is possible that some practitioners perceive that they need to maintain separate data systems to ensure they have access to the data they need for monitoring their particular interventions because they fall out of the scope of the 'standard' systems for internal record keeping (also see discussion below under 'Evaluation'). One higher-education provider interviewee spoke about capturing participant data and having "nowhere to store it". That person went on to describe how the university was building a CRM system, but it was proving challenging for the A&P practitioners to influence the design of the system sufficiently to meet their particular data needs.

Institutional records systems are widely used to monitor student success and progression activities. Almost nine out of ten (n=52) survey respondents make use of this data and it is the main source for around a third of respondents for monitoring both success (n=8) and progression (n=4) activities. Student MIS or learner analytics systems are also relatively widely used and are the main source used in half of the providers for monitoring student success and in two-fifths of providers for monitoring progression activities. The main advantages of using internal MIS/analytics systems are the coverage of the data and ease of access. A particularly useful feature of this source is security of the data, noted by over six out of ten respondents who used the student records system to monitor their student success activities. However, one advantage that MIS/analytics systems appear to have over student records systems is ease of reporting.

The current findings show that higher-education providers draw upon a range of internal and external data sources to inform their monitoring activity. Internal student data systems are perceived to be important for evidencing and nuancing the results of particular student success activities. However, our findings suggest that higher-education providers prefer 'official' datasets as this data has been externally validated and provides a clean dataset. Interviewees highlighted the advantages of working with HESA datasets, over and above internal data, as it provides the most definitive source and access to new indicators. The OfS gaps data is positively perceived as it highlights how the provider is doing in relation to the overall patterns and gaps. Working with historical datasets and the inherent time-lags presents a common barrier that can impede effective monitoring. Speeding up access to external datasets was one of the main recommendations for the OfS to emerge from the higher-education provider interviews. Dealing with multiple datasets presents a further barrier, as it is not always easy to reconcile the data across them.

Ethical considerations

There are ethical implications of data-driven interventions in student access, participation and success, including the use of learner analytics to target specific groups with different protected characteristics. These issues can be complicated by wider issues associated with disclosure (e.g. rates of disclosure of a disability) and the impact this has on the accuracy and reliability of the data. Over two-fifths of survey respondents (44%, n=24) said that disclosure issues are a barrier to making effective use of data for monitoring student success and/or progression activities and a fifth (22%, n=13) identified accuracy of data as a further barrier. The interviews highlight that the issue of disclosure and the implications that this can have for the reliability of the data sources are currently at the forefront of their minds in

relation to post-entry datasets. One higher-education provider interviewee noted that they do not currently monitor for sexual orientation because there are insufficient numbers.

Datasets for evaluation

In the latest guidance on A&P plans,¹⁷ the OfS emphasises the importance of evaluating the impact of A&P activities on learner outcomes. To support higher-education providers to develop and strengthen their current evaluation practice, the OfS has developed an evaluation self-assessment tool and standards of evaluation evidence, along with guidance on how to use them.¹⁸ This represents a step-change for many providers and there is increasing recognition among the stakeholders consulted that evaluation of A&P work is an area in need of strengthening. Some of the higher-education provider stakeholders pointed to a lack of a coherent approach to evaluation within their institution. Others, particularly small and specialist providers, reflected that, at present, data is used primarily to inform narrative reports, rather than for any systematic evaluation. In the absence of robust impact evaluation, providers rely on participant feedback or trends over time to assess whether an intervention is working effectively. The range of secondary information sources and data systems that higher-education provider survey respondents use to support the evaluation of A&P access-related, and success and progression, activities are shown in Figures 17–19 below. Perceptions of usefulness for each of the data sources and systems are also shown.

¹⁷ <https://www.officeforstudents.org.uk/publications/regulatory-notice-1-access-and-participation-plan-guidance/>

¹⁸ <https://www.officeforstudents.org.uk/publications/standards-of-evidence-and-evaluating-impact-of-outreach/>

Figure 17: Secondary information sources and data systems used to support the evaluation of the outcomes for participants who take part in A&P access to higher education-related activities, and sources rated among the top 3 (base = 73)

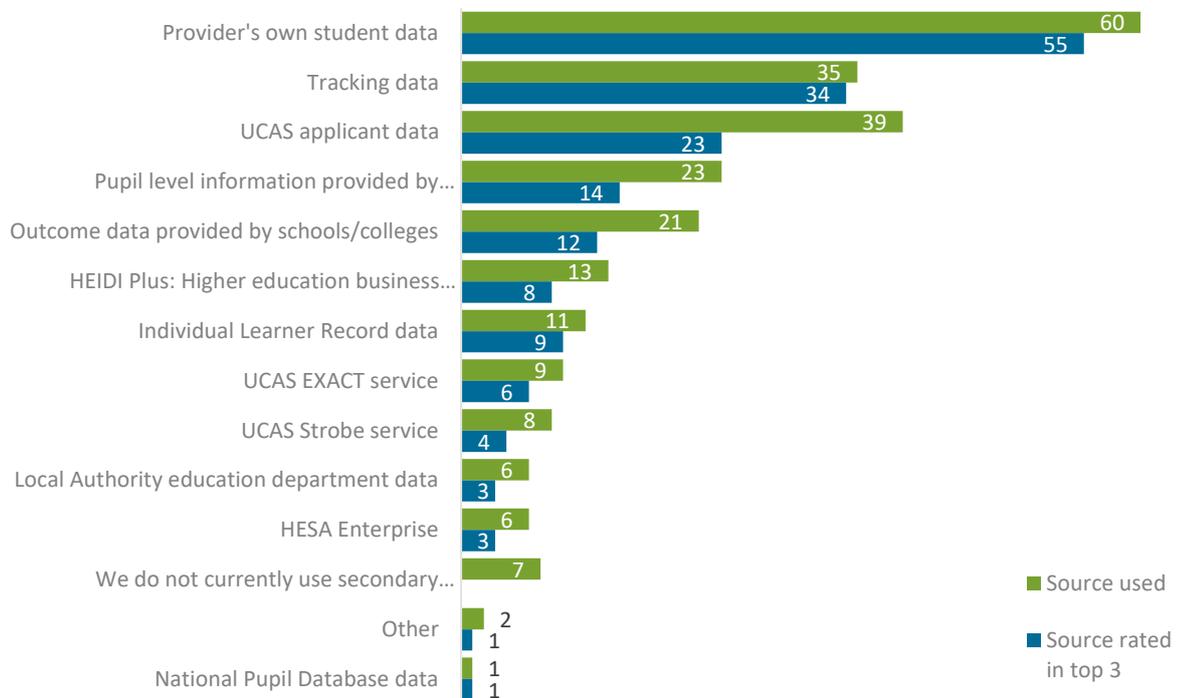


Figure 18: Secondary information sources and data systems used to support the evaluation of the outcomes for participants in A&P student success related activities and whether most important source (base = 58).

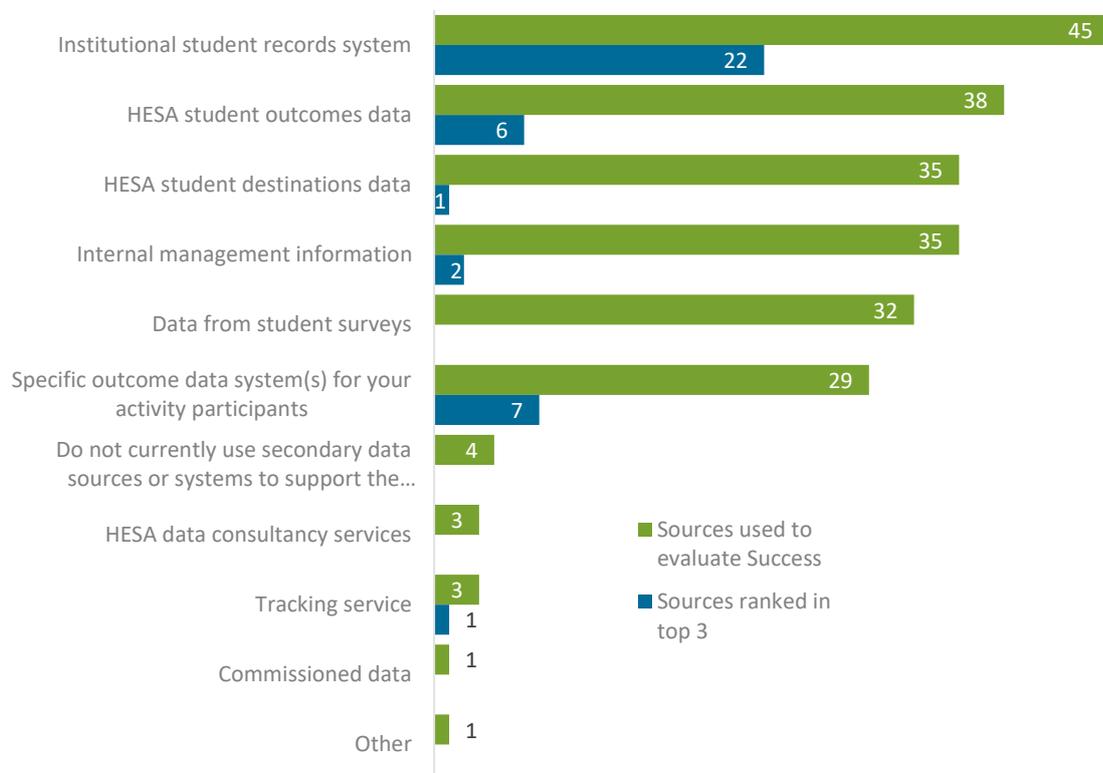
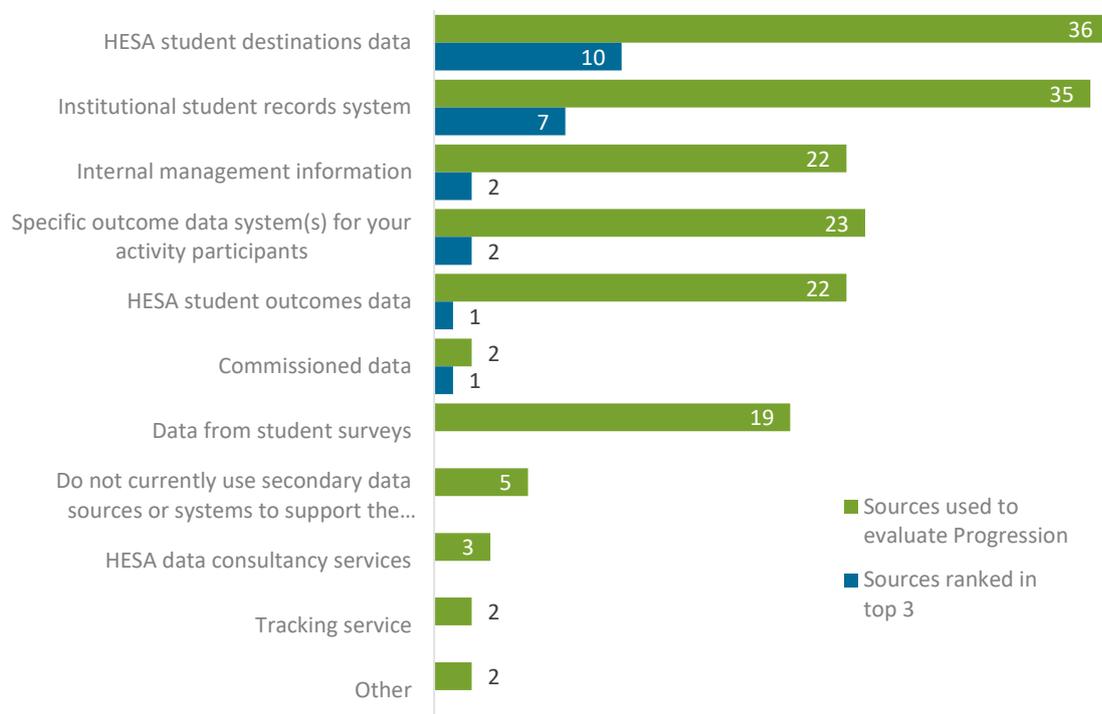


Figure 19: Secondary information sources and data systems used to support the evaluation of the outcomes for participants in A&P student progression related activities and whether most important source (base = 58).



While most providers indicated in the survey that they made use of secondary data to support the evaluation of their A&P activities, as many as one in ten (n=7) do not currently use data to evaluate their access-related activities and four survey respondents do not use it to evaluate success and progression activities. In the absence of this data, it is unlikely to be possible to control for other factors, such as learners' characteristics that could impact on outcomes, and this has implications for the robustness of the evaluation and the conclusions that can be drawn from the analysis.

The in-depth interviews revealed signs that some practitioners are developing more proactive and strategic approaches to evaluation, prompted by, and in response to, the OfS's guidance. This includes thinking about how administrative data can contribute to the evaluation of their A&P initiatives, adopting aspects of good practice in terms of identifying the data and measures to capture the results of the work appropriately, and planning in advance what datasets and systems will be needed. However, access to appropriate data sources appears to be a fairly prevalent problem in relation to both the evaluation of success and progression activities (see Chapter 4 for more details about barriers).

External data sources for evaluating A&P activities

External datasets, including linked administrative data, fulfil an important role in higher-education providers' approaches to evaluating A&P activities, particularly in relation to tracking and measuring outcomes. Tracking across educational phases is more problematic than tracking within higher education because it requires data-sharing with others.

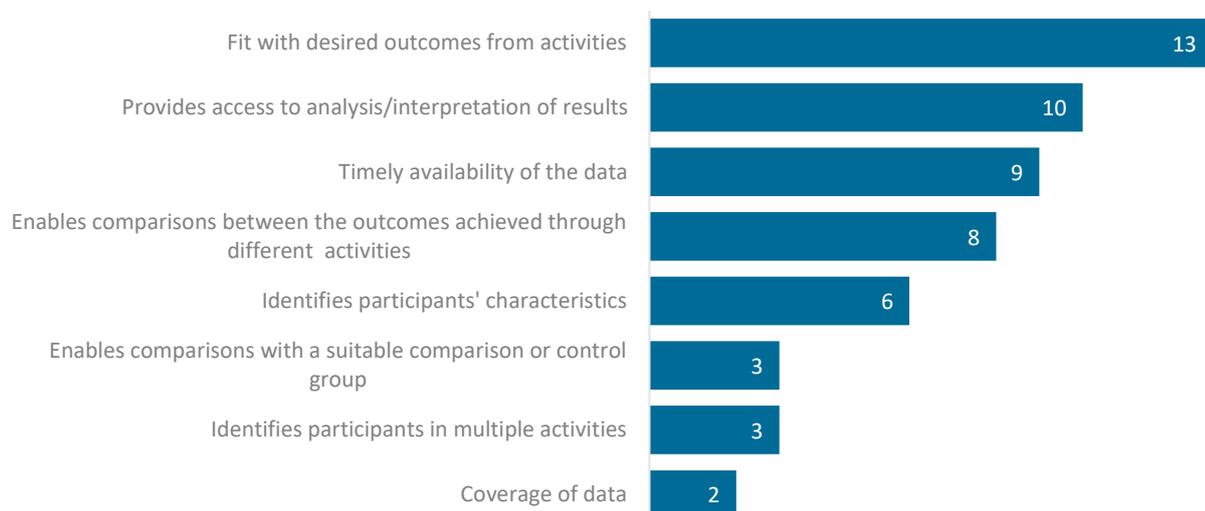
According to interviewees, key data sources on outreach participants' outcomes over time include UCAS STROBE data and data from partners, in addition to the providers' own registration data. There are several strengths and limitations associated with using the various data sources for evaluation purposes, as summarised in the sections that follow.

Evaluating outreach activities

Pupil-level information provided by schools / colleges

Just under a third (n=23) of respondents to the survey use pupil-level information provided by schools/colleges (e.g. teacher feedback) to inform the evaluation of their outreach. Three-fifths (n=14) of respondents who use this source rate it as one of their top three most useful sources. This source is perceived to be useful because the data reflects the desired outcomes and providers have access to analysed and interpreted results (Figure 20).

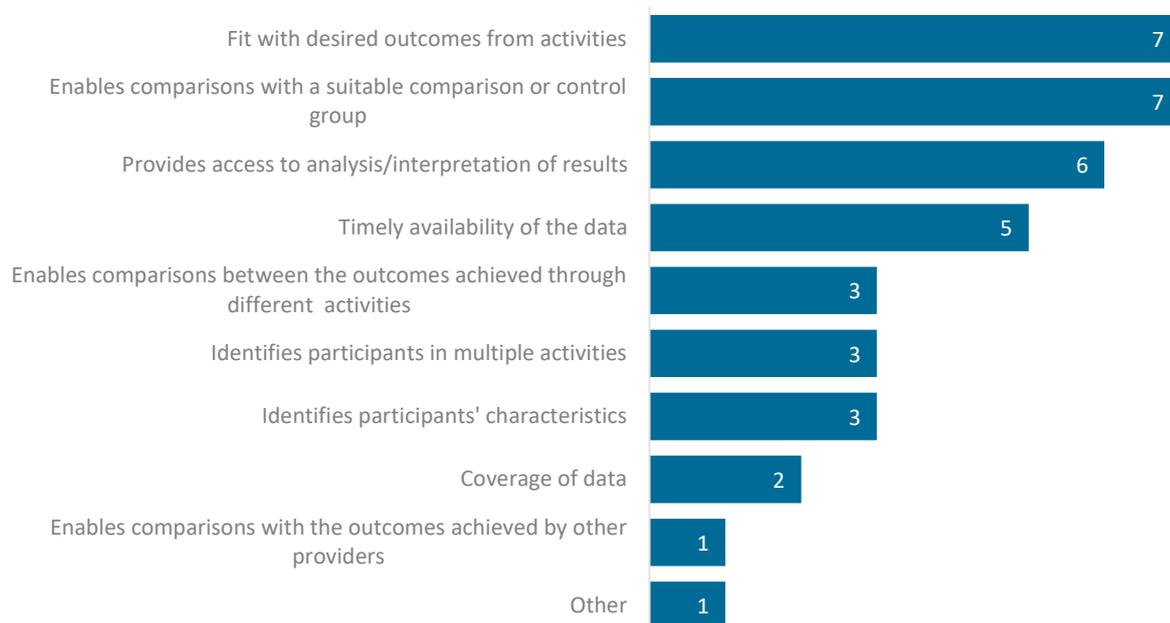
Figure 20: Reasons why pupil-level information provided by school/college is useful for supporting the evaluation of the outcomes for participants in A&P access-related activities (base = 14).



School/college-produced data

Outcomes data provided by schools/colleges is used by over a quarter (n=21) of respondents to evaluate their outreach, and three-fifths (n=8) of these respondents rated it as one of their top three most useful sources. It is perceived to be useful for capturing the desired outcomes and allowing comparisons (Figure 21).

Figure 21: Reasons why outcome data provided by schools/colleges is useful for supporting the evaluation of the outcomes for participants in A&P access-related activities (base = 11).



Interviewees suggested that sourcing data and information on outreach participants' outcomes locally could be a relatively successful approach, depending on the context. As might be expected, the nature of the relationships in place with schools stands out as being key to accessing longitudinal data through schools and colleges. In general, this approach appears to be most feasible for relatively tightly-targeted and well-established interventions that are more progressive or intensive in nature. This is because these types of interventions require higher levels of engagement with external stakeholders and there is more incentive to share the data that higher-education providers need. However, a potential issue for the A&P community is that these types of data-collection methods are unlikely to be replicable across providers and the results are unlikely to be comparable with those produced through other methods of capturing impact.

Tracking data

Using data on higher-education progression produced by the tracking services is an attractive option in relation to the evaluation of outreach activities because it provides a shared service that can facilitate access to longitudinal datasets. Just under half (n=35) of survey respondents use tracking data to inform their evaluation activity of access interventions, and all but one of these respondents (n=34) rate it as one of their top three most useful sources. In contrast, only a minority of survey respondents use tracking data to inform the evaluation of success (n=3) and progression activities (n=2).

Stakeholder interview findings suggest that data matched across different administrative datasets (e.g. NPD, ILR, and HESA) via a tracking service provides comprehensive data detailing contextual data, in addition to outcomes data. Tracking service membership can also help providers adopt a consistent approach and prevents the need to source outcomes

data independently. Survey findings show that higher-education providers perceive that tracking data provides a range of useful features to optimise evaluation processes of access interventions, including enabling comparisons between outcomes achieved, being able to identify participants who have engaged in multiple activities and the ability to identify participant characteristics (see Figure 22). Time-lags between outreach participation and the availability of datasets before higher-education entry is one limitation of tracking data. Some higher-education providers have also been hampered in their use of historical data on previous cohorts by compliance issues. However, where data is available to be matched, there are opportunities in relation to tracking outcomes at different stages of the student journey.

Figure 22: Reasons why outcome data provided by tracking services is useful for supporting the evaluation of the outcomes for participants in A&P access-related activities (base = 34).



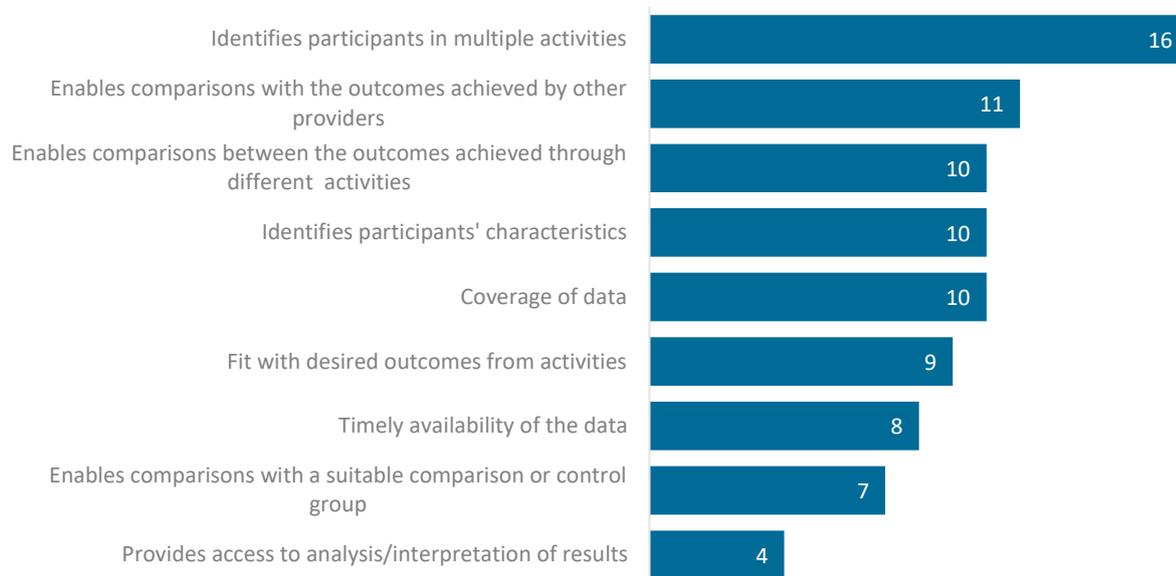
UCAS data

UCAS applicant data is used to evaluate outreach by just over half (n=39) of respondents to the survey, and nearly three-fifths (n=23) perceive it to be one of their most useful sources. UCAS also provide a range of data services which facilitate access to wider datasets across institutions. One in ten respondents (n=8) said they use the UCAS STROBE service to evaluate their access outreach, and this was perceived as one of the most useful sources for half of these respondents (n=4). A similar proportion (12%, n=9) use the UCAS EXACT service, and among this group it was one of the most useful for two-thirds (n=6).

The main reasons why UCAS data is perceived to be useful are good data coverage, facilitating comparisons with outcomes achieved by other higher-education providers, and the timeliness of the data (Figure 23). UCAS is viewed as a timelier source of data than HESA data because it is available at the end of every application cycle. However, because it is based on applications/acceptances, it is less reliable than HESA data that identifies

learners who are known to have actually progressed to, and entered, higher-education destinations.

Figure 23: Reasons why outcome data provided by UCAS applicant data is useful for supporting the evaluation of the outcomes for participants in A&P access-related activities (base = 23).



UCAS STROBE is a useful source of high-level data on higher-education applications for participant cohorts. However, some higher-education providers would like to be able to do more nuanced analysis of the outcomes data, which could be underpinned by finer-grained, if not individual-level, data. Furthermore, a couple of interviewees raised concerns about the suitability of the comparison groups used in STROBE in relation to their particular target groups and contexts. The STROBE service derives comparison data from prior attendees at UCAS open days, who are likely to be Year 12/13 learners, rather than pre-16, learners. This raises the question as to whether it is a valid approach to use these learners as a comparison group as they are more likely to already be on the trajectory of progression to higher education. Deriving a more representative comparison group to include a wider group of learners targeted by higher-education outreach activities would provide a more robust approach. UCAS is not universally used by all higher-education providers, which means there is incomplete coverage of the data. For example, Open University applicant data is excluded from UCAS, which means that mature learners are more likely to be omitted.

Evaluating success and progression activities

HESA data

HESA student *outcomes* data is used to inform the evaluation of student success activities by two-thirds (n=38) of respondents, but only a minority (n=6) perceive this as the most useful source. Fewer survey respondents (n=22) reported using HESA student outcomes data to inform the evaluation of student progression activities and this is not perceived as a

useful data source to support this activity (only one perceived it as the most useful source). The coverage of HESA student outcomes data is perceived to be the most useful aspect to support the evaluation of success activities.

A similar proportion of survey respondents said they use HESA student *destinations* data to evaluate their student success (60%, n=35) and progression (62%, n=36) activities. This data source is perceived as the most useful source of secondary data used by 44% (n=10) of survey respondents. This contrasts with only one respondent who perceived the HESA student *destinations* data as a useful data source for evaluating success activities. Coverage of the HESA student destinations data, along with the ability to compare a provider's outcomes with the outcomes of a suitable comparison group and with the outcomes achieved by other providers, are the features which make this source particularly useful for evaluating student progression activities.

Insights from the stakeholder interviews suggest that HESA data is considered a 'clean' and 'fixed' data source that enables consistent analysis to support the evaluation of success and progression-related activities. The use of national definitions and categorisation is perceived to be a further benefit to facilitate benchmarking with other providers and across the sector. Aligning HESA data with internal data systems, planning cycles due to significant time-lags, and institutional aims and objectives, were the main barriers to emerge from the stakeholder interviews. There is also a perception among smaller providers that meaningful analysis can be a challenge due to suppression techniques applied to small sample sizes.

Paid-for data

When seeking to access data for evaluation purposes, interviewees suggested that going down the route of using a paid-for data service such as HESA data consultancy services was advantageous because, despite the cost (which could be prohibitive for some smaller providers),¹⁹ the data was timelier and data providers were able to be more responsive to the needs and requirements of individual higher-education institutions. However, survey findings show that only a minority of higher-education providers currently use paid-for data (n=1–3) to support their evaluation of success and progression activities.

Student survey data

Over half (55%, n=32) of survey respondents said they used data from student surveys such as the NSS to support the evaluation of their student success activities, and a further third (n=19) used it to evaluate progression activities. However, student survey data is not perceived as the most useful source of evaluation by any survey respondents. A number of evaluation resources are available to providers, such as the financial support toolkit that enables higher-education providers to use tried and tested surveys that are then linked with secondary data (e.g. HESA). Increased engagement with existing toolkits would help higher-

¹⁹ HESA is a non-profit organisation with an open data strategy. The charges applied for consultancy services in addition to data cover the requirements to use analytical or other resources.

education providers strengthen their evaluation methodologies by ensuring valid and reliable instruments were used.

Internal data sources for evaluating A&P activities

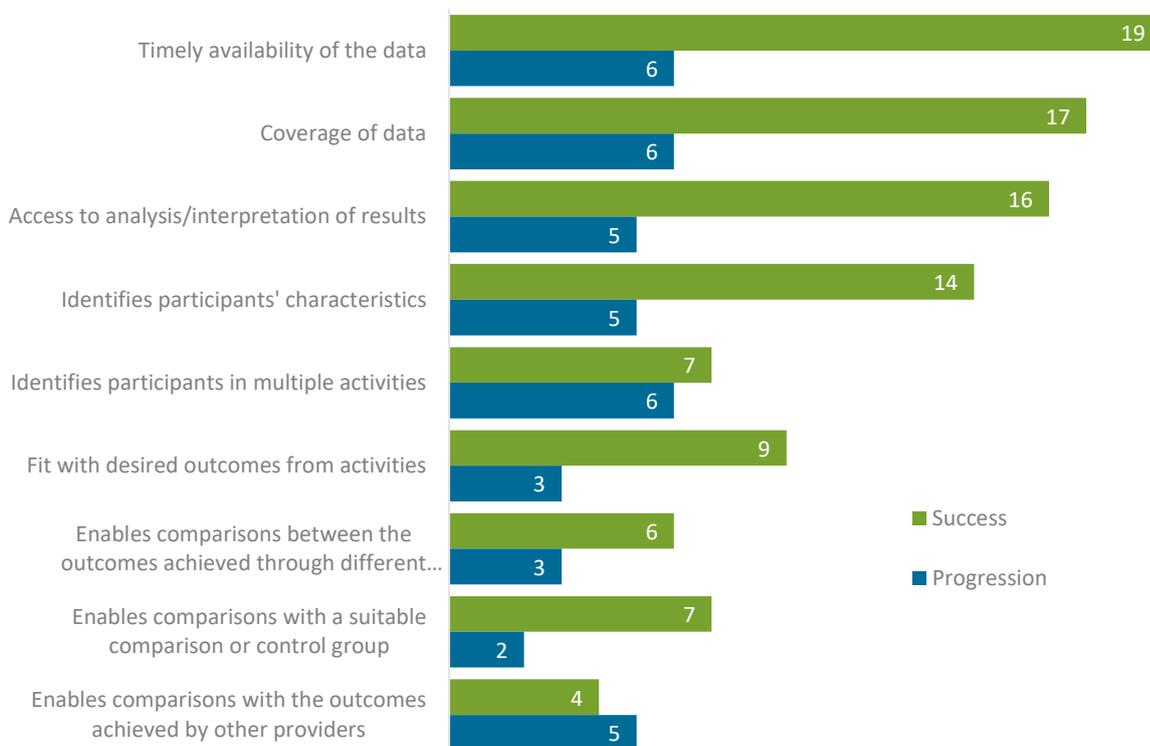
Eight out of ten (n=60) respondents to the survey said they used data already collected and held by the provider to evaluate their outreach activities. The timeliness of this data makes it particularly useful, along with a focus on the desired outcomes.

Student records systems

The institutional student records systems were commonly used by respondents (n=45) for the evaluation of success activities. This was perceived to be a more useful source of secondary data to support student success (n=22) compared with progression activities (n=7). The timeliness of the data means that student records systems are particularly useful. Data coverage, including student characteristics, and access to expertise to analyse and interpret were also perceived as key advantages of using this data source for evaluation purposes (Figure 24).

Student records system are perceived to be less useful in identifying participants in multiple student success-related activities and there is not always a good fit between the data and the desired outcomes. In contrast, student record systems are perceived as useful for comparing the outcomes of those who participate in student progression activities.

Figure 24: Reasons why institutional student record systems are useful for supporting the evaluation of the outcomes for participants in A&P success and progression activities (base = 7–22).



Stakeholder interviews identified a range of benefits for using institutional student record systems for their evaluation of student success and progression activities. Access to data on key indicators such as retention, engagement and degree outcomes was one of the main benefits to emerge.

“We have various tracking methods in place for the students who, for example, being part of some kind of intervention. For example, students who may have visited our learning and development centre, they would be tagged on our system. We might look at their attainment at the end of the year.” [INT11]

There is also the opportunity to enhance student records by adding in new data fields when institutions’ own systems are used. Using a student record system to underpin an evidence-driven approach to student success interventions can help to secure engagement from students which, in turn, makes it easier to obtain the information required. However, the level of sophistication, in terms of the types of outcomes and variables that can be brought to bear, depends on the internal systems being used, particularly the extent to which a comprehensive learner analytics framework is in place.

Bespoke outcome data systems

Specific outcome data systems are more frequently used to evaluate student success activities (n=29) compared with progression activities (n=23), although only a minority of respondents (n=2-7) perceived these systems as the most useful to support these activities. Reasons why this source was particularly useful varied, reflecting the specific nature of the bespoke data system(s) they have put in place for their activities. This approach moves beyond the use of secondary data and draws on data sourced directly from individuals.

Demonstrating impact of A&P activities

None of the institutional stakeholders who took part in the in-depth interviews were setting up their A&P delivery in a way that would allow them to implement an experimental evaluation design, although some could see the potential for this in future. It is recognised that identifying or devising suitable control or comparison groups represents a significant challenge when evaluating the impact of A&P activities for providers. Large-scale activities which are typically over-subscribed present an opportunity for individual higher-education providers to randomly assign learners to a treatment and control group for the purposes of comparison, but this is not feasible for under-subscribed activities and those targeted at smaller sub-groups. Alternative approaches are, therefore, required.

There are a number of services that can support higher-education providers in developing a more robust approach to impact evaluation, i.e. ‘Type 3’ research, as defined in the OfS *Standards of Evaluation Evidence*.²⁰ In the absence of access general population data, UCAS STROBE has derived comparison groups from existing UCAS contacts. However, interviewees who had used the UCAS STROBE service were sceptical about the methodology used to identify comparators and the lack of transparency in the methodology,

²⁰ <https://www.officeforstudents.org.uk/publications/standards-of-evidence-and-evaluating-impact-of-outreach/>

as this makes it challenging to improve or replicate future analysis. For example, one interviewee highlighted that it would be helpful to derive a comparison group that takes institutional type into account, together with the course outcomes of interest (in this case, highly-selective courses such as medicine).

HEAT has used propensity score matching (PSM) to create matched and treatment groups differentiated by the level of engagement in outreach activities (e.g. learners who have engaged in one intervention compared to more progressive schemes). To ensure that these learners have not been exposed to interventions and/or that there are no leakage effects from those who have engaged, it is necessary to draw on a wider population dataset to identify an appropriate comparison group. Developing more reliable comparison groups to demonstrate the impact of outreach activities requires access to national administrative datasets such as NPD, which is inherent with challenges in accessing the data and subsequent time-lags.

Discussion and future models

Availability of datasets for targeting

Access to reliable and timely data for targeting is a prevalent concern for providers, particularly in the context of outreach. Measures with a high degree of validity and reliability are required to ensure outreach reaches those who would otherwise not be on a pathway to higher education. The multiplicity of potential target groups, and the inter-sectional nature of educational disadvantage, means it is unlikely that a single dataset or indicator is sufficient. All of the current indicators have inherent weaknesses in terms of reliability and/or utility; the challenges become even greater when coupled with lack of data expertise at the provider level and scope for misuse. Guidance for providers on indicators and measures, the circumstances they can be applied, and the inferences that can be drawn, is required.

School-level analysis based on NPD sourced through the tracking services is being used as an alternative to published data, which can be hard to manipulate and interpret, or area-based measures which are unreliable, for targeting schools for participation in outreach. Currently, this type of data is only available through the trackers, and access is hampered by the requirement for separate applications to be made in line with NPD-linked data protocols. Expediting the availability of this data, through facilitating access to NPD on an ongoing basis, would be desirable, and wider dissemination of this type of existing tool for targeting would be important.

The development of new datasets for targeting is an area where the regulator can play a leading role, and we understand developments are already underway, led by DfE, to publish a school-level progression measure. However, providers are likely to require support to operationalise the new data source as a tool for targeting. There is scope for developing and piloting a new resource for school-level targeting (including new DfE measures and matched NPD data for schools), with the objective of agreeing new sector-wide, shared resources/lists for targeting.

Lack of availability of datasets for targeting outreach activities for mature and part-time learners emerges as a major gap. The OfS should explore options for a mature learner planning dataset (possibly drawing on the HEFCE Adult Higher Education (AHE) measure²¹). This could involve establishing a working group of higher-education providers with a focus on mature and part-time learners with a view to making recommendations for new data and measure(s) to help providers identify and prioritise mature learner disadvantage.

Availability of datasets for targeting individuals

There is a sense that more work is needed to identify who the most appropriate target groups are, especially for participation in sustained and progressive programmes of activity. There is a lot of interest in using FSM indicators from providers, but many find this type of data difficult to obtain. Furthermore, there are problems when using this data for individual-level targeting, although it is a useful indicator of the relative levels of educational disadvantage across schools (or as part of a suite of contextual information). Measures that take account of pupil-level progression through compulsory education have a key role to play (e.g. thinking about groups with above average KS2 who are falling back, absent or disengaged by KS3). However, the challenge of recording these factors and the complex nature of attainment patterns means these types of measures are probably best operationalised through partnerships and dialogue between higher-education providers and teachers in schools, rather than through a strict application of data based on prioritisation rules (as we are proposing in section 7 on school-higher education links). This approach could be backed up by a Working Group, involving schools and higher education. The groups would be tasked with developing and testing a methodology for using the pupil-level datasets to prioritise learners for inclusion in outreach, drawing on the expertise of the OfS's analytics team to identify groups at risk of not fulfilling their higher-education potential, and taking account of existing approaches to composite measures and data triangulation.

The recent publication by the OfS of *Association between characteristics of students (ABCS)* provides new and experimental access and continuation measures.²² The aim of the ABCS set of analyses is to support a better understanding of how outcomes vary for groups of students who have different sets of characteristics. This should be a useful resource, enabling higher-education providers to target students who are most likely to benefit from specific types of outreach interventions and/or success and progression activities.

While the above approaches will mainly benefit outreach practitioners in terms of targeting, they will also increase the availability of measures that can be applied to monitor participation. They also offer a basis for further benchmarking and the consistent analysis of targeting outreach activities within local areas/regions.

²¹ See, for example, Crawford & Greaves (2013) A comparison of commonly-used socioeconomic indicators: their relationship to educational disadvantage and relevance to Teach First. IFS Report R79. Institute for Fiscal Studies. <https://www.ifs.org.uk/comms/r79.pdf>

²² <https://www.officeforstudents.org.uk/publications/associations-between-characteristics-of-students/>

Datasets for monitoring

There is no single common measure used to assess improvements in equal opportunities. Current approaches usually involve a basket of indicators/measures and there is a trend towards using composite indicators. There is a tension between the desirability of having a relatively small number of measures which facilitate benchmarking and reporting on progress nationally, and the need to use a multiplicity of measures in order to capture the complexity of the issues and contexts which affect student success at the local level. Higher-education providers favour flexibility to use measures that best capture disadvantage within their particular context.

There is a gap between how institutions currently use administrative data for monitoring and optimal use. The need to work within multiple systems and a lack of data expertise are key barriers. In an ideal world, practitioners would like to be able to use one joined-up dataset that links individual learner records across the lifecycle. When asked about the prospects for increasing the effectiveness of data use, many of the stakeholders consulted highlighted the need for more timely and ready access to the NPD, ILR and HESA datasets at low, or no, cost.

The tracking services are seeking to provide a joined-up approach to monitor participation (i.e. drawing on pupil-level characteristics such as home postcode profile and prior attainment, with their outreach involvement, recruitment activities, subsequent higher-education applications and enrolments, and related continuation, attainment and progression outcomes). Recording of participation in activities in a systematic way can also be used to monitor activities and provide outcomes data to inform evaluation that takes account of variation in student characteristics. Integrated systems increase the potential for sharing aggregated data locally and nationally. The use of dashboards provided by the trackers ensures monitoring data is transparent to stakeholders. There is also scope for increasing joint arrangements with schools for sharing data, recording and monitoring, as discussed in Chapter 6.

Use of linked datasets for tracking and evaluation

The ability to link data between educational phases is a requirement for effective outreach evaluation. Tracking services play a key role in linking data between educational phases for the purposes of tracking outcomes, although one can envisage a range of different approaches to tracking arrangements (see discussion in Chapter 3). In the future, providers would like access to datasets that link pre-entry and post-entry data, for example, tracking service data linked to student and graduate outcomes. This could be achieved via a Data Lab or data request service, or through the pooling of data centrally for the purposes of data linking and evaluation.

Demonstrating outcomes for participation in A&P interventions relative to a comparison or control group, in order to understand impact, represents a key challenge. There is a lack of understanding about the range of suitable methodologies that should be applied, and the importance of adopting a proportionate approach to evaluation methods, among practitioners. Support to develop suitable methodologies for analysing outcomes data from A&P interventions is fundamental to making progress on the effective use of linked

datasets. Part of the challenge will be to establish whether shared, ‘standardised’ methodologies could be applied across a suite of different interventions, or whether highly individualised, activity-specific methodologies would be most appropriate. Each of these routes would have different implications in terms of the future models and mechanisms for effective data use.

Development of ‘standard’ methods of working with linked data for evaluation purposes increases the scope for data controllers to make available agreed data extracts for these purposes. The Ministry of Justice Data Lab is an example of where this type of approach has been successfully implemented in another policy area (see [Appendix 2](#) for further details and examples from other sectors). The Data Lab acts as an intermediary for the purposes of data linking, based on common agreement about the sharing of data between practitioners and data controllers for evaluation purposes.

Within the existing higher-education data landscape, the Financial Support evaluation toolkit is a model which could be adapted and used to evaluate other types of intervention. The strength of this approach is that it offers a pragmatic solution which takes account of the level of analysis expertise within higher-education providers. It also draws on specified data definitions and access to data through HESA, the designated data body. Expanding this model also has the potential to enhance the capacity for data use and evaluation among practitioners.

03. Engagement with tracking organisations and value for money

Here, we explore higher-education providers' engagement with the three tracking services. The perceived costs and benefits of current arrangements and implications for effective working are reviewed, moving beyond a purely monetary assessment of value.

Introduction

Three tracking organisations currently operate on behalf of higher-education providers to offer services that support targeting, monitoring and evaluation of A&P outreach activity. Tracking organisations have their origins in local widening participation partnerships, emerging as a legacy of Aimhigher. They operate on a subscription basis. HEAT, the tracking organisation with the largest membership, has also benefited from funding from HEFCE to support its expansion and roll-out. The respective services have slightly different funding and governance arrangements. Overall, they represent a 'bottom-up' solution to the challenges of targeting, monitoring and evaluating outreach, by pooling resources and providing expertise and systems to support data use. Minimising duplication and maximising economies of scale are the main benefits of membership of a tracking organisation. Other 'bought-in' systems exist to support data capture and evaluation, including approaches being developed by ImpactEd²³ and the Careers & Enterprise Company.²⁴ These currently offer solutions for addressing particular activity-related research and evaluation questions, rather than the types of cross-activity and multi-institutional systems data that the tracking services have sought to put in place.

Operating context and services provided by trackers

There are similarities in terms of the system capabilities and types of support provided the three tracking services, but there are also a number of differences that can impact on the ways in which higher-education providers engage with their tracking service. The key similarities and differences are summarised in Table 2. A more detailed review of the functionality of the three tracking services is provided in [Appendix 3](#).

²³ <https://impacted.org.uk/>

²⁴ <https://tools.careersandenterprise.co.uk/>

Table 2: Summary of the characteristics of the three tracking organisations

Tracker Feature	Similar approach	Different approach
Online database solutions providing a shared repository for data and reporting tools using 'bottom-up' systems ²⁵	✓	
Datasets to enable individual/institution targeting	✓	
Survey functionality to capture feedback	✓	
Linking to external datasets (HESA common to all 3 trackers) ²⁶	✓	
Provision of activity and participant reports via dashboards	✓	
Access to aggregated data from members	✓	
Provision of data expertise and analysis	✓	
Provision of support and networking		✓
Funding and governance arrangements		✓
Database functionality		✓
Approach to legal gateway and role of data controllers and processors		✓
Evaluation support		✓

Differences in **database functionality** is one of the main distinguishing features between the three trackers. For example, the HEAT database includes additional functionality designed to support the administration of events and outreach-related activities such as work with student ambassadors. HEAT and EMWPREP offer facilities that enable members to securely share information and tools designed to support outreach teams. HEAT provides support and networking opportunities via regular Members' Forums, Data Users Network meetings and Research Group workshops. EMWPREP provides support and networking opportunities via newsletters, operational groups and standard and bespoke training sessions for partners. Other types of support offered by all three trackers are more ad hoc and locally based.

There are also differences between the **reporting functions** of the three trackers. AWM and EMWPREP coordinate evaluation and research across members. For example, AWM has developed standardised evaluation toolkits and methodologies (e.g., pre/post and quasi-experimental methods) for use by all members to measure the short-term outcomes

²⁵ Databases have been developed collaboratively with members and are capable of interfacing with university data systems (i.e. through API) in order to minimise the administrative burden

²⁶ The HEAT Track approach provides the most comprehensive arrangements currently for ongoing longitudinal tracking in terms of the range of administrative datasets that are accessed to provide data showing their educational outcomes.

as well as the medium- to longer-term impacts of their activities. This centrally-managed approach to research, evaluation and reporting ensures consistency across the partnership, which helps to facilitate comparisons and benchmarking; it also reduces burden on individual providers and duplication of effort, which helps to ensure value for money. EMWPREP undertakes analysis and produces monitoring reports bi-annually (interim and end of year) on behalf of partners, which contribute to institutional-level monitoring and reporting. In contrast with AWM, EMWPREP designs and implements bespoke methodologies to support higher-education providers' evaluation activity. In addition to their core offer, all three trackers have sought to be responsive to members, and some aspects of the services have been developed to support the objectives of particular providers, albeit with agreement from the members via a democratic approach to decision making.

Despite differences between the trackers, there is a high degree of cross-over in terms of the types of data the systems collect and store. The services have recently collaborated in order to map the coverage of members' outreach. Data was pooled in order to identify schools and colleges across England that engaged in at least one outreach activity during the academic years 2016/17 and 2017/18.²⁷ Through this mapping project, the tracking services have agreed a methodology to enable outreach coverage to be mapped in the future. An ongoing issue for tracking services is how best to measure the intensity, type and duration of the activities. Through the NCOP, tracking services are collaborating to derive a common classification framework, which will ensure greater consistency in terms of these measures in the future and, therefore, enable comparisons.

Current membership of tracking services

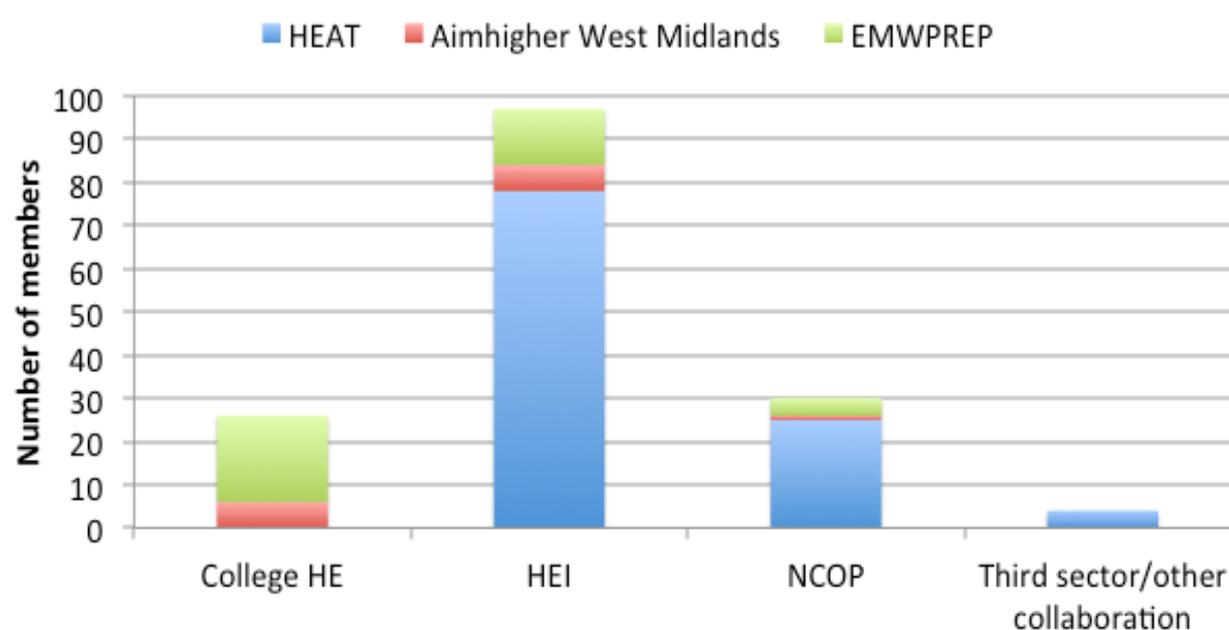
Tracking service membership exceeds 150 (with a small number of providers involved with more than one service). The membership primarily comprises universities and colleges, with a small number of third-sector organisations (see Table 3). HEAT benefited from HEFCE Catalyst Funding, which supported the expansion of the service to new members. As such, HEAT has largest number of members overall, including the third-sector providers. EMWPREP and AWM operate on a smaller scale and are relatively localised. Their differentiated membership model also includes colleges (see Figure 27). NCOP partnerships are required to subscribe to a tracking organisation by the OfS as a condition of their funding, to enable them to monitor engagement in NCOP-funded activities (which are additional to institution-specific activities, which may also be tracked through membership of a service).

²⁷ Mapping output available at: <http://www.emwprep.ac.uk/emwprep-news-may-2019-issue-13/>

Table 3: Members of tracking services (April 2019).

	1 tracker	2 trackers	All	% of total
Higher-education institution	88	5	93	61%
College of higher education	26		26	17%
NCOP partnership	29		29	20%
Third sector / other collaboration	4		4	2%
Total	14	5	152	100%

Figure 25: Members by type and service²⁸



Larger providers, defined in terms of level of investment in A&P, make up the largest proportion of tracking service members, and the share of the membership increases significantly when the level of A&P investment is £2,000,000²⁹ (see Table 4). Collectively, the trackers have national reach, but membership is more highly concentrated in the East and West Midlands, South East and Greater London – the regions where the tracking services originate from (see Figure 28).

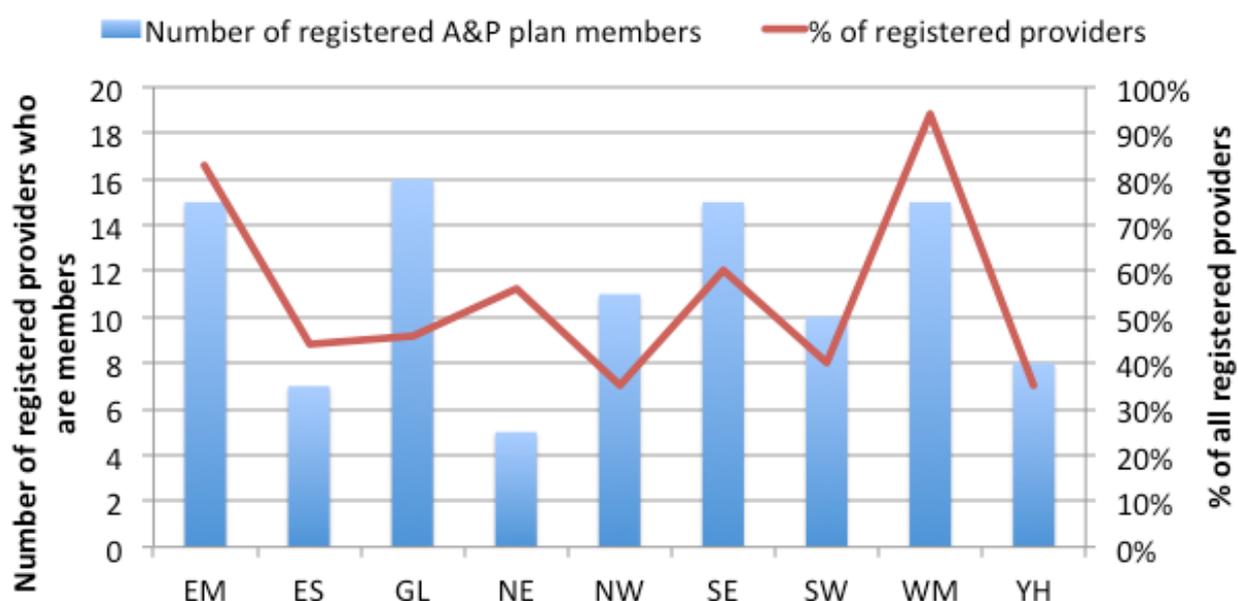
²⁸ A target for HEAT as a result of HEFCE funding was to increase the number of English member HEIs to at least 80 English by May 2017 (there were 107 members in April 2019 of which 78 were HEIs).

²⁹ The 198 providers with an agreed APP for 2019-20 had annual budgets for A&P activities (excluding financial support) ranging from £11,000 to well over £19 million. Although the average works out at over £4 million, nearly two-fifths of providers (38%) had a level of investment below £500,000.

Table 4. Membership by A&P Investment in 2019-20³⁰

£,000	Member	Non-member	All	% of total
Up to 150	4	32	36	11%
150-500	5	29	34	15%
500-2,000	8	23	31	26%
2,000-6,000	27	7	34	79%
6,000-10,000	26	5	31	84%
>10,000	32	-	32	100%
Not included in A&P dataset	17	-	-	-
Total	119	97	198	51%

Figure 26. Members by region (2019-20)



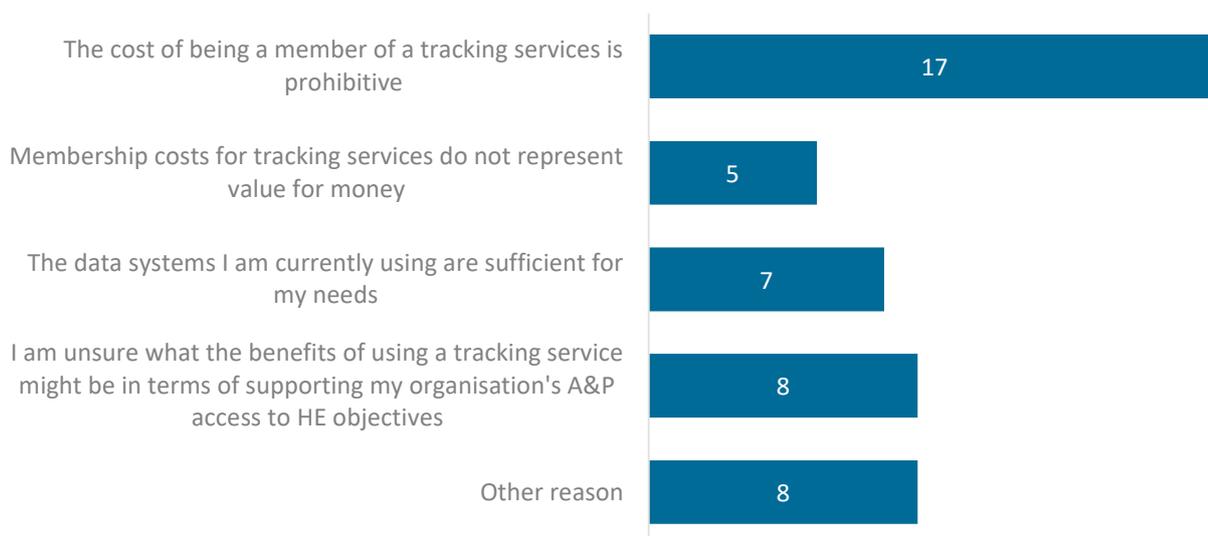
Over half (n=40) of survey respondents are members of tracking organisations. Most are large higher-education providers (n= 23), defined in terms of A&P spend, reflecting the national trend. A total of 28 respondents reported that their institutions do not currently use a tracking system. While 10 of these indicated that their institution was definitely not considering using a tracking service, 15 said that they were considering membership of

³⁰ We note that the OfS register lists 229 providers as having an access and participation plan in place for 2019-20, however the figures are based on the 198 given in the A&P plan dataset at December 2018. The OfS register lists all the 352 English higher education providers officially registered by the OfS at this date. <https://www.officeforstudents.org.uk/advice-and-guidance/the-register/the-ofs-register/>

HEAT.³¹ The remaining respondents did not know whether their institution was considering membership of a tracking service.

It appears that overall cost, rather than perceptions of value for money, is the principal deterrent (see Figure 27). More detailed information given to higher-education providers about the functionality and capabilities of tracking services could encourage some of them to join (as well as help current members to maximise the benefits of their membership), as some non-members said they were unsure as to how membership would support the achievement of their A&P objectives. A similar number felt that their current systems were sufficient for their needs. A range of other reasons for not having membership were cited, including delays in the availability of data and being new to A&P and the use of data. A lack of data on some groups, such as mature students, in the mainstream datasets is identified as a limitation and key barrier to effective data use (see Chapter 4). The findings suggest that a lack of functionality in this respect could also deter providers with high proportions of students in these groups from using a tracking service (see Figure 27).

Figure 27: Barriers to membership of a tracking service (base = 28)



Use of tracking services to support targeting, monitoring and evaluation

The stakeholder consultation highlights that the level of engagement with tracking services can vary over time as a result of changes in priorities, or turnover of personnel. The consultation suggests that some higher-education providers do not always have the capacity to make full use of the tracking service, or lack the necessary expertise in data processing, analysis and interpretation. One interviewee reported that their institution had been a member of HEAT for two years but it had not been able to get the full benefits because of the investment of time required to implement it and use it. Where capacity and expertise

³¹ AWM is not open to new members.

are in place, however, there is evidence that the system is fully embedded and providers are using it to record all their outreach-activity delivery.

Our survey findings suggest that higher-education providers engage with tracking services in different ways,³² which are likely to reflect the context within which the provider operates, its A&P objectives and, as noted above, its capacity and expertise in data use. Of the 40 respondents who are members of a tracking service, 36 are members of HEAT. The top four services used by members across the trackers are: systems to monitor access/outreach activities (n=33), training and support to use the system (n=31), and data for targeting outreach on areas, schools or colleges (n=30) and member working groups/forums/events (n=28) (see Figure 28, on the next page).

Data prioritisation and targeting

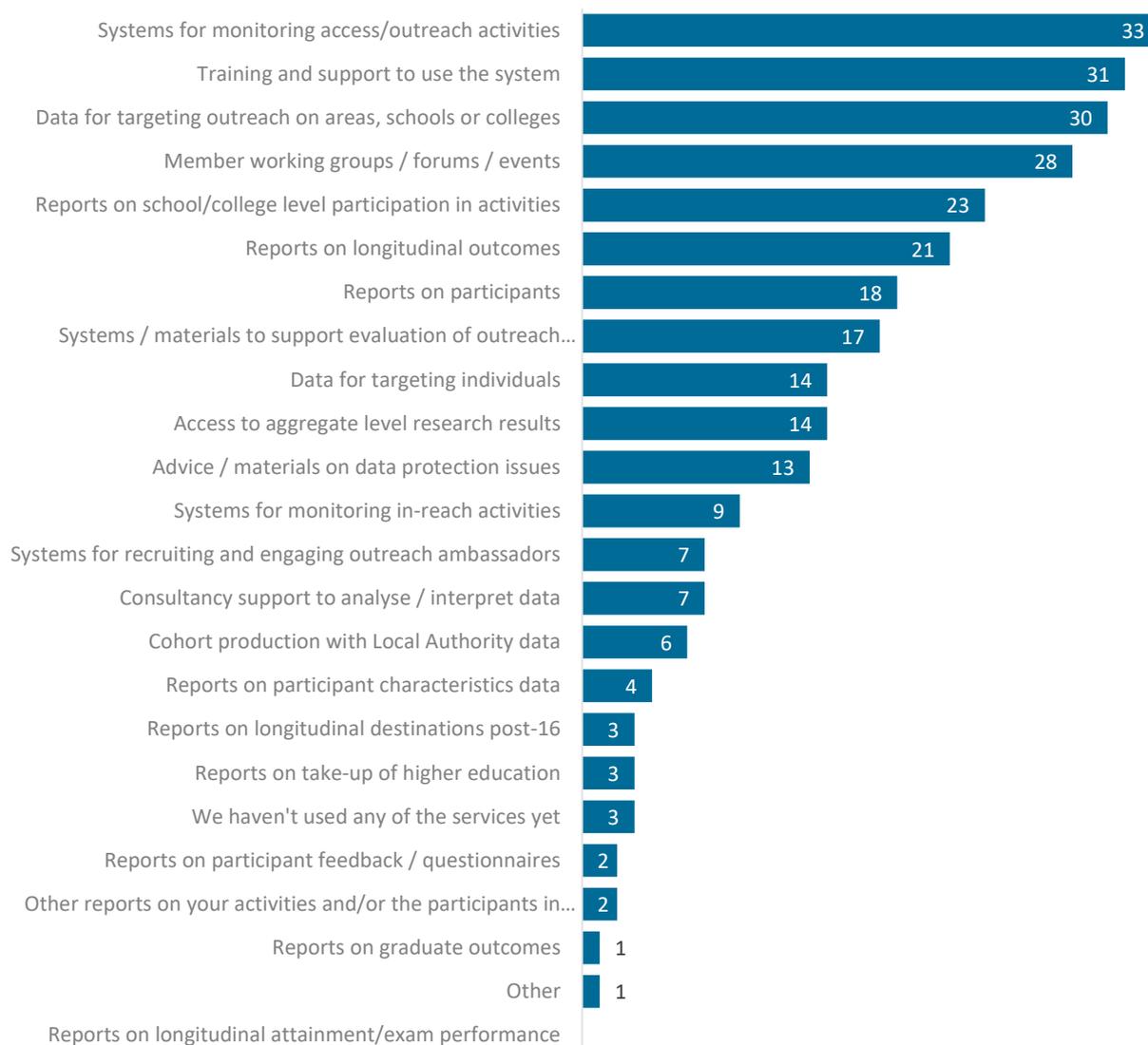
As illustrated above, the majority of higher-education providers, irrespective of which tracking service they are a member of, use the service to target and prioritise A&P outreach at an *area level*. Drawing on school and college-level data, merged data on deprivation indicators and POLAR from tracking services as an alternative to the published datasets was perceived to be a benefit of membership by some interviewees. Directly accessing and analysing schools' data is time consuming; accessing this data via a tracking service is perceived to be more efficient. In addition, merged data on deprivation and POLAR would not otherwise be available to higher-education providers. Accessing school characteristics and performance data was described by one provider as their “biggest time sink”.

A minority of higher-education providers access data to target individuals via trackers, but to a lesser extent than area-level targeting (n=14). EMWPREP members also reported using local authority data to produce a cohort list and HEAT members draw upon data and look-ups to verify individual-level data.

Only a minority of survey respondents engage with a tracking service to inform targeting of A&P success-related activities (n=8), which supports the view that trackers are predominantly used at the access stage in the student lifecycle.

³² No survey data was available for how HEPs engage with AMWM tracking service

Figure 28: Higher education providers' engagement with tracking services (base = 40)



Systems for monitoring

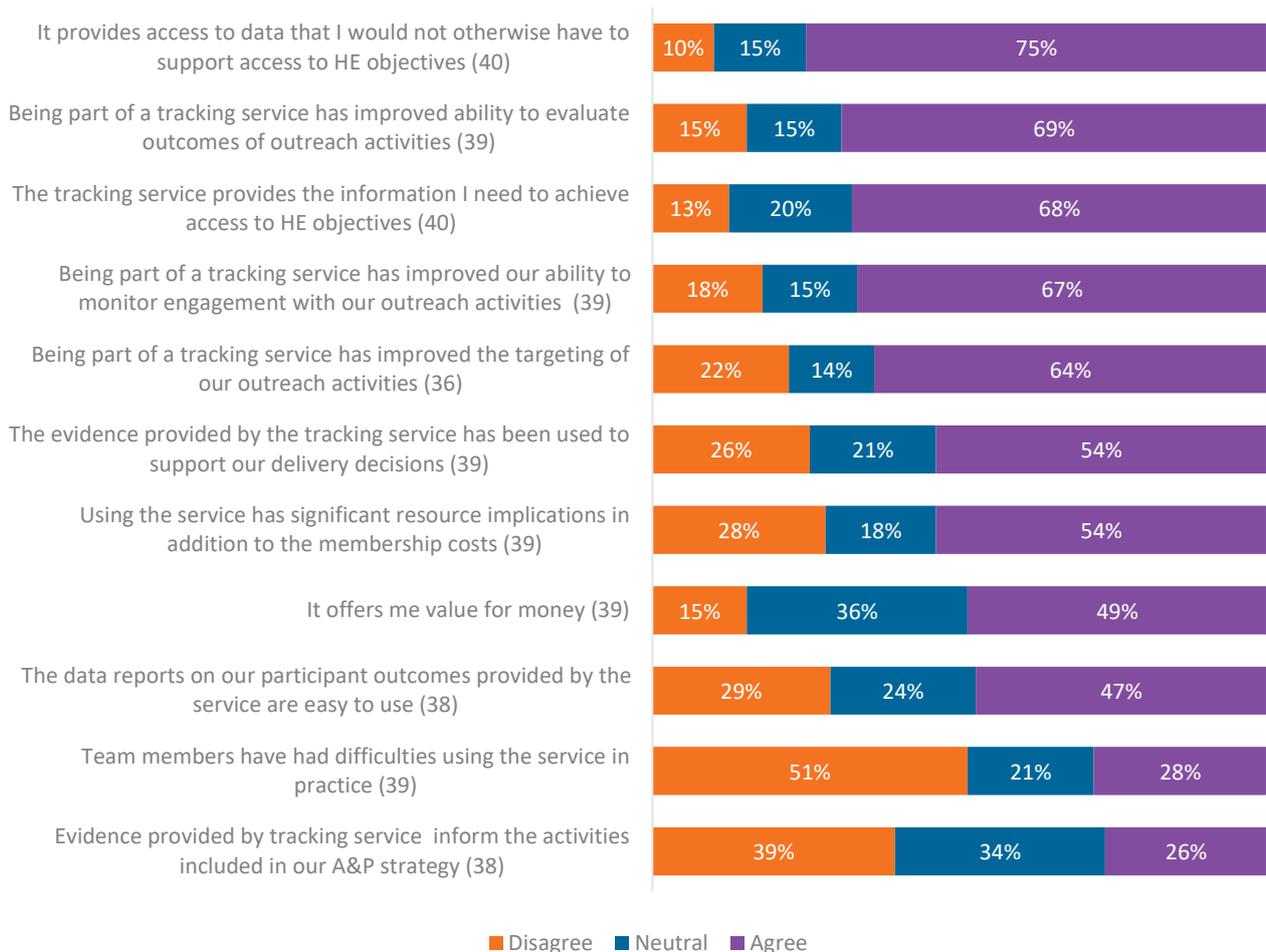
Systems for *monitoring* access/outreach activities were the most commonly-used aspect of tracking services. Nearly a quarter of HEAT members (n=9) reported using the service to monitor in-reach activities. Only a small number of survey respondents (n=4) said they used tracking services to monitor student success and progression activities and only one respondent perceived their tracking service to be the most important source for monitoring their progression-related activities.

The key advantages of using a tracking service for the purposes of monitoring identified in the depth interviews included the assurance that the systems in place for data capture are GDPR compliant. The provision of a secure data system that can be linked to other existing sources was perceived to be a further advantage of using a tracker for monitoring activity.

Evaluation activity

A total of 17 tracking service members use it to access systems and materials to support the evaluation of their *outreach* activities; a small proportion (n=3) also use a tracking service to support their evaluation of success and progression activities. One in seven members perceive that being part of the tracking service has improved their organisation’s ability to undertake evaluation (see Figure 29). As noted above, the systems and materials available to members varies between the trackers, with some offering more bespoke services and others a more standardised approach. The nature of the service on offer could help to explain the variation in usage, depending upon the requirements of the higher-education provider. A lack of robust tracking data stands out as the main reason why members are not taking advantage of the services to inform their evaluations. This is because outreach participants often have to be tracked over a number of years before outcomes can be captured. Delays in the data becoming available, due to data processing, and time taken to secure access to it, compound the problem of data lags.

Figure 29: Views on tracking services



Using findings from reports on participants and longitudinal outcomes can support evaluation activity. Our survey findings show that 22 of the 40 tracking service members use participant reports for this purpose. A further 21 (all HEAT members) use reports on

longitudinal outcomes. The findings from the depth interviews suggest that longitudinal outcomes data is being used to benchmark progress against other providers and institutional key performance indicators (KPIs), rather than inform evaluation activity. For example, one provider interviewed has KPIs based on higher-education destinations data from HEAT. This can be beneficial for providers where only a minority of participants who participate in the outreach activities they provide progress to their institution. Further research could seek to explore the extent to which tracking service reports inform higher-education providers' evaluation activity.

The cost of tracking

Tracking services have different funding models and provide services that are not directly comparable. The current annual cost for providers ranges from around £4.5k to £10.5k, and the costs are associated with different types of service in terms of database developments, access to data and mechanisms of support, including evaluation, as discussed above. For NCOP partnerships, the costs of membership of a track service was based on a proportion of their total budget (approximately 0.2 per cent of a partnership's allocation, increased to 0.65% for HEAT in Phase 2).

The costs of tracking membership differ according to provider context. For example, one interviewee from a large provider described the membership fee as “nothing for an institution like ours”. In contrast, a small provider reported cost was prohibitive, reflecting the message conveyed as part of the OfS consultation that smaller providers perceive the costs of a tracking service to be disproportionately high. Membership level and type also make a difference. One provider that is intending to move from using a CRM to HEAT indicated it was worth the money, whereas another who used a CRM to feed into a broader CRM across the lifecycle did not.

Tracking members incur additional costs to operationalise services, which can vary according to provider characteristics. Depth interview findings highlighted the extent of this variability. For example, one provider said their capacity was a 0.6 FTE, whose role was to run the database, plus a full-time person with responsibility for inputting and analysis. Another estimated staffing costs of approximately £70k per year for data and evaluation tasks. The costs of data analysis can be significant and are not always accounted for. For example, approximately 20% of the lead practitioner's time for each activity is spent on data and evaluation-related activities, according to another provider.

The stakeholder consultation highlighted that many higher-education providers found it challenging to quantify the amount of resource required for the data side of their A&P work. There was a tendency to focus on data entry and management, rather than capabilities (such as statistical analyses), infrastructure, executive oversight, investment in academic/practitioner research, and staffing costs for data collection. Our findings suggest that this is an under-resourced area that requires a culture shift to ensure sufficient expertise is provided in the future.

Value for money

Despite half the survey respondents reporting that they had experienced difficulties using the tracking services, most perceive that the trackers deliver a range of benefits which support effective targeting, monitoring and evaluation of A&P, and just under half agree that the service offers value for money (see Figure 29). Perceptions of value for money from the stakeholder consultation largely depended on whether the provider considered use of the service to be optimised, and whether the potential benefits in terms of longitudinal tracking had been realised. A range of experiences were captured through the stakeholder consultation, with one provider reporting that it was still waiting to receive its first tracking report, and another reflecting that they would be focusing on how to use the service more for the administration of their activities, and to drive data collection and reporting in future, to maximise the benefits of membership.

Tracking service members have to pay on an annual basis, which potentially creates some tension in the initial stages because higher-education providers may have to wait beyond the first few years before the tracking data on their participants becomes available (depending on the age of participants and their distance from higher-education entry). Storing up data and then entering it onto the system when the participants are at the point of appearing in tracking data is identified as one strategy to reduce costs and maximise benefits in the first instance. Another provider mentioned that there had been internal pushback on joining a tracking service because the service did not deliver a broader function across the student journey. None of the tracking services currently offer whole-system solutions for all stages of A&P activity in the student lifecycle.

Although it is hard to put a precise figure on the resources currently used in relation to the tracking services, insights from the research suggest that the level of direct and indirect resource is considerable. The estimated costs associated with different data systems, compared with a tracking service, are shown in Table 5. The potential benefits that can be derived as a result of this investment are summarised in Table 6 in order to begin to develop a cost-benefit framework.

Table 5: Estimated cost of tracking services and alternatives.

Considerations	Estimate of Tracking service costs	Cost of alternatives
Systems	<ul style="list-style-type: none"> Tracking service (£4k-£10,500 pa) 	<ul style="list-style-type: none"> CRM system/software license (£60-£240 pm) Database development (£50-£250 per day)
People	<ul style="list-style-type: none"> Data analyst (£20-25k pa) Data entry (£8-20 p/h) (or upload from parallel system using API) 	<ul style="list-style-type: none"> Data manager (£30k per annum) Data analyst (£20-25k per annum) Data entry (£8-20 per hour) Data protection officer (£35k+ per annum)
Datasets	<ul style="list-style-type: none"> Tracking service planning datasets (in subs) Linked data through tracking service (in subs) 	<ul style="list-style-type: none"> Published datasets (£0) Pupil data (£0+) Data analysis (estimated at two weeks work for data analyst) higher education Applicant data (£0-£2,000) Higher Education Student data (£600-£5650) ACORN (£7k-£9k p/a) Resources involved in sourcing matched data (estimated at weeks work for data manager)
Expertise	<ul style="list-style-type: none"> Support on using the system, data analysis and networking/sharing 	<ul style="list-style-type: none"> Database training/consultancy (£500+ per day) IT consultancy (£70-£110 p/h) Research consultancy (£500+ per day)

Table 6: Benefits of using a tracking service (*Aspects where there is additional added value above what HEPs would otherwise put in place in the absence of tracking services).

Considerations	Benefits of tracking services (when optimised)
Data	<ul style="list-style-type: none"> – Provision of planning data enabling higher-level targeting – Look-ups and matched data enables individual-level targeting – Access to unpublished metrics assists targeting activity* – Enables tailored targeting approaches, dialogue with schools* – Provides learner outcomes data*
Systems	<ul style="list-style-type: none"> – Enables centralised data capture (cross-institutional)* – Enables improved data capture (e.g. frequency of interaction/contact hours and first age of engagement)* – Provides a safe data repository with security measures in place – Provides effective data management processes
Analysis/Reporting	<ul style="list-style-type: none"> – Provides ‘Live’ data dashboards enabling effective monitoring – Activity reports to measure progress against KPIs and benchmarking – Provision of multi-activity reports on individuals to inform the evidence base for sustained interventions over time (e.g. as part of progression framework approach)* – Cross-sector outcomes provided (outcomes in different educational phases and beyond own institution)* – Aggregate datasets assist with benchmarking and contributes to strengthening the evidence base*
Support & expertise	<ul style="list-style-type: none"> – Enables proficient set-up and ongoing training of staff (face-to-face, online and resource-based) – Networking enables collaborative working among members* – Support for reporting and analysis ensures effective allocation of resources – Evaluation planning contributes to strengthening the evidence base

Collective approaches to data use and sharing have obvious potential for efficiency and cost savings, through splitting costs, minimising the duplication of effort and pooling resources, support and expertise on data issues. As noted previously, the time saved as a result of sourcing data from tracking services was one of the key benefits to emerge from our stakeholder consultation. **Time savings** can be achieved even in relation to easily-accessed datasets (e.g. configuring data from published datasets to prioritise outreach interventions).

Enabling access to data which higher-education providers would not otherwise be able to obtain is a further key benefit. Access to NPD data and the ability to develop a demographic profile of state-funded secondary schools in England using Pupil Census data (i.e. home domicile data for each pupil in a school is used to provide a summary of the

disadvantage profile for the school) was highlighted as particularly beneficial by interviewees. Data look-ups (e.g. postcode checkers) is a further function that enables proficient working practices, including identification of specific gaps and tighter targeting.

Benefits in enhancing access to, and sharing the costs of, data are likely to be gained in relation to **data linking across the student lifecycle** for longitudinal tracking purposes. This is enabling providers to gain an enhanced understanding of the relationship between outreach provision and the attainment and progression of the young people who engage with it. This, in turn, helps higher-education providers to adopt an evidence-based approach to planning outreach delivery. Using data from a tracking service also enables providers to gain a better understanding of how interventions are addressing aims and objectives and closely monitor delivery against operational plans. Considerable knowledge and expertise is required for these aspects, which higher-education providers rarely have in-house (with the exception of larger, selective providers in some instances). HESA and UCAS apply charges for data consultancy. Processing work on very small samples sizes may not be cost-effective; however, economies of scale can be achieved if large data requests are submitted because the costs do not increase exponentially. Higher-education providers recognised that tracking service membership provides efficiencies by bypassing various agencies, and establishing data-sharing arrangements themselves.

Non-monetary benefits of the tracking services emerged in our findings, including access to **training and support** and the development of alternative methods and techniques. Expertise in data and analysis has been provided by the tracking services in a range of ways through formal and informal support, information and training. Although the focus is outreach related, support is being drawn upon across the student lifecycle to inform post-entry work. In terms of training and support to use the system, 31 of the 40 tracking service member respondents to the survey said they had used those services. A similar proportion had taken part in member working groups/forums. A minority of members (n=7) also access consultancy services to support analysis and data interpretation; with more (n=13) getting advice on data protection (see Figure 28).

There is evidence that tracking services are **pooling methods** to support higher-education providers with data analysis and interpretation. The HEAT groups approach and AWM offer ways of comparing outcomes for cohorts of students taking account of attainment levels and other characteristics. This enables a tailored approach for the analysis of outcomes achieved that moves beyond whole-class results. Comparing different cohorts with similar characteristics can be useful for benchmarking and assessing the effectiveness of targeting, although this is not a substitute for the more sophisticated statistical techniques that attribute impact by comparing treatment and control groups. However, our findings suggest that this is not yet being widely used. Value-for-money considerations also need to account for the social benefits, including whether outreach and higher-education provision is becoming more diverse and better meeting student needs as a result of the services provided.

In the absence of tracking services, it is likely that higher-education providers would find alternative solutions. Some providers consulted were using a bespoke CRM to store data and were expanding internal information systems for pre-entry tracking (SITS or Tribal

EBS). However, evidence suggests that higher-education providers would struggle to replicate tracking services in several areas: 1) being able to associate students with multiple activities; 2) analysing engagement across a particular institution; 3) providing an overview of activity engagement across multiple institution; and 4) tracking enrolments into different institutions.

Furthermore, feedback from the OfS consultation identified the pros and cons of the different services and an overall perception that practitioners would not welcome a situation whereby higher-education providers are required to (a) use a tracking service and/or (b) a particular tracking service. As discussed throughout this report, higher-education providers have different needs which would not be addressed by a one-size-fits-all approach. A differentiated system enables higher-education providers to engage with a service that suits their requirements. A service that has built considerable local knowledge and expertise, and achieved buy-in from local partners, is able to draw on this to offer value for money.

Return on investment

Sector-wide tracking is a key aspect of the added value of tracking services, and aggregated data can be used by providers for benchmarking. There is also a sense that the aggregated results provided enable an enhanced sector-wide view of A&P outreach delivery that can inform the evidence base on the effectiveness of different interventions. Seeing ‘*the big picture*’, as well as the opportunity to share ideas, is a further key return on the investment. There is still, however, some way to go in optimising the amount of activity data included in the tracking services, and putting in place effective, more consistent, categorisations of types of activities.

Respondents to the OfS consultation perceive that the OfS has a role in ensuring tracking processes serve the needs of the higher-education sector, for example, by highlighting ‘cold spots’ or overlaps in provision. Our interviews also highlighted that more could be done to share activity data both locally and nationally to provide outreach practitioners with a coherent overview of A&P interventions on the ground. Tracking data represents a partial picture of the outreach provision that is taking place as the tracking services do not include all A&P providers in England. Therefore, mapping based on data via the tracking services cannot be used to identify those individual schools or colleges that have not taken part in any outreach. Developing a more robust way of pinpointing the ‘cold spots’ in outreach delivery would be useful in addressing overall coverage issues.

Discussion and future models

Improvements to the current model

There appear to be efficiency gains to be made from providers collaborating and sharing the costs of accessing data. However, whether these can be realised in practice depends on local circumstances. Furthermore, some providers are currently paying for services that are not being used to full effect; optimising the use of the services currently available would increase the return on the investment of these higher-education providers. The varied level of engagement across the higher-education sector is a key weakness of the current approach, in terms of realising the benefits of centralised outreach data systems. Our findings suggest that a range of factors impact on the extent to which members are using the full functionality of systems. This includes: 1) contextual factors including institutional size and characteristics of the student population, A&P objectives, APP spend and infrastructure; 2) time and resources, including staff capacity and expertise; 3) difficulties with GDPR compliance; and 4) cost of membership.

Whether one defines value for money at the provider level, or in terms of the costs and benefits to the sector, ultimately the calculation relies upon there being a positive outcome in terms of provider outreach and the achievement of policy priorities for the access and progression of under-represented groups in higher education. The challenge is realising the value of data-related aspects to promote equal opportunities in targeting, monitoring and evaluation activity. Introducing frameworks that promote collaboration and a consistent approach could help to improve efficiencies and enhance the sector-wide benefits of tracking organisations. This could also provide an opportunity for further 'bottom-up' solutions to be developed. Common aspects of the framework include: datasets for planning and prioritising activities; look-ups for individual-level targeting; reporting functionality for monitoring; central databases for recording and storing data securely; and access to matched data in respect of this for longitudinal tracking. The effectiveness of the current systems could be improved if there were further development in these key areas:

- Enhancing engagement in the systems by clarifying expectations in terms of internal staffing and resourcing requirements, and ensuring resources are effectively implemented. **The OfS has a potential role in endorsing the legitimacy of staff time spent on evaluation-related aspects.**
- Maximising data collection and tracking capability by clarifying the legal gateway and agreeing data responsibilities in terms of data control and processing. **Clarification from the OfS on the use of public task will help to expedite data collection.**
- Building the capability to analyse, interpret and use data within higher-education providers and enhancing the use of analysis and reports at different levels (to further influence decision making and promote buy-in to the systems). **The OfS has a potential role in clarifying regular reporting requirements, as well as providing support to develop standardised analysis methodologies.**
- Developing a set of consistent validated measures (or evaluation questions based on a relevant theory of change) that could be used to compare practice nationally. **There is scope for trackers to assist partners to assess the implications of the**

results for their practice through further developments to create ‘standard’ methodologies for analysing tracking data.

- Expediting and streamlining the processes for accessing and reducing the costs associated with access to administrative datasets, particularly for targeting (i.e. NPD). Data for outcomes tracking has been relatively easy to access, via HESA, although there are costs involved.
- Opening up opportunities for tracking membership to a wider range of providers, through different levels of service being offered by existing organisations or new intermediaries that focus on core offer/outputs.

There have been questions about whether different systems would help to improve the current practice. Having different systems means there is some duplication of effort, but the benefit of this is that the offers and levels of service suit different types of higher-education provider and are a forum for providers to collaborate to generate ‘bottom-up’ solutions where applicable.

Feedback from higher-education providers and the tracking organisations suggests that having a range of options available is the preferred route, reflecting the diversity of the sector and the difficulty of defining any ‘one-size-fits-all’ solution that takes account of the wide range of institutional contexts and priorities. Tracking services have particular strengths, as well as providing some common features. Depending on their offer, some of the services are supporting member-led developments around evaluation in a range of ways. In addition, research and consultancy services are being developed. For example, EMWPREP provides in-depth bespoke research services to partners, negotiated with them annually to take forward particular research and evaluation. AWM has developed an evaluation toolkit and HEAT is extending support on methods to include an evaluation planning toolkit in the future.

Future models

Table 7 (below) considers the opportunities for strengthening the arrangements for using linked administrative data in the future. It also considers different directions of travel that could support further efficiencies and added value. Further centralising data linking and reporting activities for learner outcomes is a potential option, either as part of an OfS function or via a third party ‘Data Lab’ function, as exists in other policy areas.

There is also scope to move towards a more centralised approach, to support the evaluation of higher-education outreach at both the national and local level. In the past, the use of administrative data for evaluation purposes has been a centralised function. For example, HEFCE tracked and analysed participant outcomes in the Aimhigher Summer School Programme.³³ Contextual factors were taken into account by linking data to NPD.³⁴ The sharing of common objectives and ‘standardised’ types of delivery in terms of the data consistencies are key strengths of this approach. However, there would be resource and

³³ <https://data.gov.uk/dataset/961fea6e-2f7d-46f5-910e-38367f6ba8b2/aimhigher-summer-schools-data>

³⁴ <https://data.gov.uk/dataset/d842096b-19d0-439b-b728-9df9d9b6ce99/summer-schools-evaluation-2012-strand-3-matched-pupil-data>

logistical issues involved. While the extent and quality of data collected by outreach providers has improved, data collection and collation is not sufficiently consistent to enable comparative data analysis of the programmes of delivery at a national level at this stage. Gaining greater consistency and agreeing standards for data collection should also be considered as a long-term aspiration, in order to support further improvements to sector-wide analysis. In developing a more centralised approach, however, it is also important not to lose sight of the value of more bespoke, localised approaches to evaluation which reflect local priorities for access and participation and contribute to the achievement of wider objectives. As with tracking service provision, maintaining a degree of flexibility to tailor approaches to local circumstances, alongside the development of centralised approaches to facilitate comparisons and benchmarking at a national level, is likely to be appropriate in the context of such a diverse sector.

Table 7: Different approaches to data linking.

	Strengths	Weaknesses	Opportunities/Potential developments
1. 'Business as usual'	<ul style="list-style-type: none"> Repository for systematic logging of data collectively Avoiding duplication/reinventing the wheel Sharing what is known about learners Secure systems Sharing the costs of systems and processes for data matching and analysis Costs are amortised across the membership Reporting is held centrally Aggregated samples (potential for richer analysis, building evidence of impact) Support/expertise from team(s) of experts Drawing on a community of support/expertise 	<ul style="list-style-type: none"> Differing levels of use/engagement Members lacking capacity to engage fully Turnover of staff/loss of expertise in members' teams The data that is available is not always used Subscriptions need to be paid over many years GDPR requirements a blocker to holding/sharing (limits use of historical data) Low response rates for consent Time delays and resources required in accessing data Fixed pricing may penalize smaller providers with fewer resources disproportionately Duplication of effort across the three trackers (although sharing also exists) Member led development processes - parts of functionality that only work in the context of particular members 	<ul style="list-style-type: none"> More consistency of data and approaches based on sector-wide guidance on the legal basis for processing Increased use of standard reporting templates Further developing the services role in facilitating a community of WP researchers working together to build the evidence base An opportunity for members to shape future research using aggregate data (including collaboration on thematic research projects) Wider dissemination of analysis and reports beyond the membership including to policy audiences Guidance and direction on GDPR issues.
2. Institutions could rely on accessing their own linked data	<ul style="list-style-type: none"> Keeps the data sharing processes 'in-house' 	<ul style="list-style-type: none"> Represents a significant burden on providers in terms of expertise, time and resources which it seems unlikely they would be able to sustain Providers lacking level of technical expertise required to work with linked data Implies multiple requests to data providers and therefore increases the burden on the system 	<ul style="list-style-type: none"> Opportunity to build up expertise of some practitioners on this issue?

Strengths		Weaknesses	Opportunities/Potential developments
3. Frameworks to support tracking organisation to develop a 'standard' core offer to support aggregation of datasets with clear legal gateway in place	As above, plus: More consistent data collection for comparative analysis Benefits for sector-wide monitoring and reporting of outcomes (while retaining scope for local flexibility to meet provider needs)	As above, depending on the incentives/directives for providers to address these issues	Expediting access to administrative data for tracking purposes through more clearly defined procedures with data providers Opening up the legal gateway could support regional or sub-regional target setting
4. Promotion of one tracker as preferred service	3. As above, plus: Widening out potential for analysis of sector-wide activities	Driven by national priorities rather than in response the service to members' needs Likely to limit provider choice and reduce flexibility for locally driven solutions Existing services not appropriate for all providers (e.g. part-time and mature learners) Probably an unpopular option	Standardised approaches to monitoring of activities and reporting on outcomes
5. Setting up new arrangement for pooling data for purposes of administrative data matching on outcomes (OfS or third party)	Development of a shared resource resources/avoid duplication of effort Expedite data access Clarity on legal gateway Processes and procedures which maximize the data linking processes	Would require readjustment of existing resources and there would be ongoing resource implications	Clarity on the national requirements for data collection Consistent or standardised approaches to reporting and tracking

	Strengths	Weaknesses	Opportunities/Potential developments
6. OfS data analyst role in evaluation using aggregated, 'pooled' A&P dataset	<p>Range of contextual data and various outcome measures taken into account</p> <p>Potential for working with most up-to-date datasets</p>	<p>Legal basis: would need to be instructed in law or permitted for management</p> <p>Applicant checks (enquiry and application process) for individual-level data</p> <p>Should not be a replacement for 'internal' teams of data experts within A&P providers</p>	<p>Would require working with different stakeholders: Mutual interest in having administrative datasets for targeting, monitoring and evaluation</p> <p>Would need to think of common definitions and classifications</p> <p>IT solutions would be essential</p> <p>Secure systems for transmitting data needed</p> <p>Need to assemble a 'person spine' to use and understand interactions across interventions</p> <p>Would require joint work on assembling data, checking coverage, data sharing agreements, common definitions and classifications</p>
7. Data providers or OfS take on providing outcomes data – data lab function drawing on NPD linked data	<p>Opportunity to link data up and down the lifecycle (to contextualise the results)</p> <p>Opportunity to create matched comparisons to facilitate estimates of impact</p> <p>Data Lab would ensure that the relevant security and legal data sharing standards are being adhered to</p> <p>Opening up linked data sector-wide</p>	<p>Ready access to pupil-level NPD data would be needed, probably via a change to the DfE privacy notice to ensure data subjects are aware of how their data is being used</p>	<p>Would need to be based on standardized agreed methodology for data linking and matching</p> <p>Secure systems for transmitting data</p> <p>Team of experts</p>

<p>8. Data providers play key role – toolkit approach to applying HESA outcomes data</p>	<p>Potential for more consistency of approaches to access, success and progression</p> <p>Basis for comparison of results based on application of standard methodologies</p> <p>Building capacity within providers for data analysis and evaluation</p> <p>Opening up linked data sector-wide</p> <p>Building on existing approach (financial support toolkit)</p>	<p>May be more appropriate to some types of interventions as opposed to others (methodologies have not been proved in relation to the suite of interventions)</p> <p>Ready access to student data would be needed, probably via a change to the HESA privacy notice to ensure data subjects are aware of how their data is being used</p>	<p>Would require development and piloting of standard methodology</p>
--	--	---	---

04. Practice-related issues

This chapter discusses barriers and solutions to effective data use, focusing on the context for practices underpinning the use of administrative data for targeting, monitoring and evaluation.

Introduction

Effective practice in using data for targeting, monitoring and evaluation of A&P activities is underpinned by factors specific to the provider context in which the data is being used. The high-level barriers are discussed here, along with potential solutions. A more detailed breakdown of the barriers experienced for targeting, monitoring and evaluation of access and success and progression activities can be found in [Appendix 4](#).

Barriers and solutions

Staffing and capacity

Resourcing issues are one of the key challenges faced by some higher-education providers that are perceived as preventing them from optimising the use of data in A&P work. Lack of staff time stands out as one of the main barriers to using data for A&P targeting access and success and progression activities. A lack of staff time is also perceived to be a prominent barrier for evaluating access activities:

Table 8: Proportion of respondents who report that ‘lack of staff time’ is a barrier to effective use of data.

Purpose	Focusing on	% perceive barrier:
Targeting	Access activities	57%
	Success and progression activities	59%
Monitoring	Access activities	53%
	Success / progression activities	41% / 41%
Evaluation	Access activities	71%
	Success / progression activities	49% / 47%

The majority (84%) of small providers reported that a lack of staff time was a barrier to targeting, while a smaller proportion of medium-sized (56%) and larger (28%) providers perceived this as a barrier. Non-members of a tracking service were also slightly more likely to identify staff time as a barrier than members, although this

probably reflects the fact that it is larger providers who are mainly members. Interview findings reflect the survey results, with numerous comments alluding to a lack of capacity, such as:

“We're not a small university, but we're also not a very large one. We do not benefit from economies of scale. That puts a real pressure on us.”
[INT11]

“There is a big cost to it and capacity issues within our management information team because there are so many draws on everyone’s time.”
[INT21]

Solutions to staffing and capacity barriers offered included whether the OfS could play a role in enabling more resources to be brought to bear in institutions to support effective data use, and mobilising more financial support. Lines of responsibility appear to be a barrier in effectively using data. Developing research protocols in relation to A&P activities that highlight which datasets are required, together with clearly defined roles and responsibilities, can help overcome these barriers.

Data and analysis expertise

Our findings suggest that getting data from participants (73%) and gaining access to appropriate datasets (67%) are perceived to be the two largest barriers experienced for targeting access activities. A lack of benchmarked data is perceived to impede the evaluation of access activities (63%). *Poor data coverage* appears to impact on effective targeting of access activities to a greater extent than monitoring and evaluation activities:

Table 9: Proportion of respondents who report that ‘poor data coverage’ is a barrier to effective use of data.

Purpose	Focusing on	% perceive barrier:
Targeting	Access activities	45%
	Access activities	31%
Monitoring	Success activities	32%
	Progression activities	32%
Evaluation	Access activities	43%
	Success activities	28%
	progression activities	44%

The results also highlight skills gaps at different levels, in terms of understanding the data, as well as technical skills in handling and analysing data. *A lack of knowledge of data sources and systems* that can be used to support targeting of A&P activities was a barrier, and this was more evident for evaluation activities.

Table 10: Proportion of respondents who report that ‘a lack of knowledge of data sources and systems’ is a barrier to effective use of data.

Purpose	Focusing on	% perceive barrier:
Targeting	Access activities	36%
	Success and progression activities	32%
Monitoring	Access activities	42%
	Success activities	36%
	Progression activities	40%
Evaluation	Access activities	42%
	Success activities	51%
	progression activities	47%

Survey respondents (52%) from small providers perceived that a lack of knowledge of how data sources could be used to support targeting of A&P access activities as more of a barrier compared with larger providers (16%).

Interview findings suggest that some data teams are exploring ways to effectively communicate data to practitioners (who tend not to be data specialists). Prioritisation models that use bandings/categories, and traffic light coding have been introduced to operationalise data for use on the ground by non-experts. Sharing expertise, facilitated by an integrated team working across the whole student lifecycle, is one way in which one interviewee’s provider is seeking to upskill staff. This ensures the benefits of data expertise required for success and progression activities are maintained throughout the various stages.

Tensions can arise when A&P practitioners are required to work with complex datasets that might be outside their existing knowledge and expertise. The local context of culture of higher-education providers is also an important consideration, as there is evidence of resistance to data-led processes from some academic staff teams. This is likely to change once the practice of using management information becomes more widespread and embedded.

Interpreting data analysis is also linked to issues around expertise, which frequently arise due to linked data files arriving as ‘flat’ files that require data manipulation and analytical skills in order to make sense of local outcome data. *A lack of expertise for data analysis and interpretation* is perceived to be a more prominent barrier for evaluation activities, with just under half (46%) of survey respondents perceiving this as a barrier for access activities and around two-thirds seeing it as a barrier to the evaluation of success and progression activities. Once again, smaller providers (68%) perceived a lack of expertise in data analysis and interpretation to a greater extent compared with larger providers (12%). Over twice as many non-members (56%) of a tracking service perceived a lack of expertise in data analysis and interpretation as a barrier compared to members of a tracking service (25%). This

suggests that tracking membership offers support in understanding data and carrying out relevant analysis. Specific training requirements to enable outreach teams to undertake more sophisticated statistical analysis (e.g. regression analysis and between-group comparisons) would provide further evaluation support as highlighted below:

“Lots of people are trying to recruit data and evaluation specialists at the moment, that’s required. I think what’s also required is a greater level of data-literacy among the broader staff-body ...be able to do some analysis themselves and be able to see an Excel spreadsheet, or table of data, and not panic, and be able to read it.” [INT16]

Recruitment issues were also raised regarding analytical expertise. For example, one interviewee spoke about the different types of skills and attributes required to be able to work with the data in this field:

“In a lot of universities, managers don’t quite understand how you have to get data analysts to really understand why they’re doing something... and the space and platform to communicate to the highest level in the organisation. We often view data people as introverted, maths-y people who stay in the cupboard and somebody else presents their results. We’ve managed to find people who do both and who are part of the endeavour, rather than just a side part. They’re central.” [INT20]

Partnership working

Effective partnership working is important, especially for data sourced from partners. Collecting individual-level data frequently requires either direct access to participants/learners or close collaboration with other intermediaries. This is especially true in the context of higher-education outreach, where schools, colleges, other providers, or other stakeholders such as parents, may provide data rather than learners themselves. The benefits of effective partnership working for sourcing data is highlighted in the example below:

“Often we’re reliant on some of our school or college partners, or the local authority, providing us with that data. One of the easier ways is on one of our progression pathway programmes ... the reason that works very well is that we have really long-standing relationships with the schools and colleges, so they’re really happy to share data with us... If we’re trying to widen that or create new interventions, partners aren’t so forthcoming in giving us data.” [INT11]

Collaborative partnership working also enables effective negotiation of data-sharing agreements, which tend to be complex and often require negotiation. Trust and a shared vision for interventions need to be put in place, not only for data to be shared and used effectively, but also so that targeted participants are actively engaged and encouraged (rather than feeling stigmatised).

Systems issues

Systems that enable effective data collection, management and analysis (e.g. tracking services) are important and it is clear from our findings that a range of systems are being consecutively used (especially for outreach) and that some providers lack such systems. As mentioned above, our findings show that a lack of knowledge about systems is a more prominent barrier for effectively monitoring A&P activities, compared to targeting and evaluation activities. Some providers appear to be using a range of systems consecutively, especially for outreach. Others have sophisticated systems for storing and analysing their own student data, particularly where the provider has moved in the direction of using learning analytics, which have the potential to bring in data from different institutional systems. Depth interview findings pointed to a lack of mechanisms for systematic and joined-up data analysis for some higher-education providers, as highlighted in the example below:

“Our challenge is that we capture initial data on people, but we’ve got nowhere to store it... I don’t believe that any university has got it absolutely nailed in terms of how they get all their data working across all their platforms in a really efficient way. We’re way more reactionary as opposed to being proactive enough” [INT14]

Smaller providers more frequently face challenges in influencing the systems used by larger providers. In particular, it can be difficult for colleges to adapt to OfS requirements.

Cost of data and systems

The *cost of data* is perceived by survey respondents to be more of a barrier for targeting, monitoring and evaluation access activities compared to success and progression activities:

Table 11: Proportion of respondents who report that ‘cost of data’ is a barrier to effective use of data.

Purpose	Focusing on	% perceive barrier:
Targeting	Access activities	37%
	Success and progression activities	22%
Monitoring	Access activities	53%
	Success activities	32%
	Progression activities	33%
Evaluation	Access activities	49%
	Success activities	27%
	Progression activities	28%

Smaller providers perceive costs as more of a barrier compared to larger providers. For example, nearly half of small providers (48%) perceive costs of data as barriers compared to 28% of large providers for targeting access activities.

Data systems which have in-built reporting tools were clearly favoured by those in operational A&P roles. The use of data ‘dashboards’ is increasing although, overall, the higher-education sector might lag behind others in terms of the data architecture. Comments from a couple of institutional stakeholder interviews point to recent developments towards rationalisation of the datasets and how they are analysed and interpreted within the institutional structures:

“My ambition for it is to accelerate this whole thing, and use more sophisticated data tools. User interfaces are going to be quite important for some of this data.” [INT20]

Discussion and future models

Clearly the solution to more effective use of data for targeting, monitoring and evaluation of A&P activities lies not just in establishing access to relevant data, the systems and processes, but also other essentials such as adequate resourcing and access to appropriate expertise, good data governance and systems/infrastructure, and effective partnerships for data.

The increasing focus by the regulator on the strategic context for evaluation within provider institutions may go some way to supporting more staffing resource for evaluation, and the research suggests that many providers have started to identify ways to address internal gaps in expertise. At the same time, findings from the fieldwork suggest that providers favour sector-wide approaches to improving data literacy, because this has advantages in promoting more consistency and exchange of ideas. A role for TASO was suggested here.

The issue of skills and expertise in higher-education providers points to the important role that sector-wide training and guidance on data use could play to promote more effective approaches. This was highlighted in comments from stakeholders who argued for development of new CPD for practitioners. Suggestions for different types of support to build capacity within institutions to use data emerged during the research. The two main areas of interest were in relation to:

- Training materials and short courses, for non-experts, to explain different A&P datasets and how they can be used for targeting, monitoring and evaluation. These could be at a high level and/or focused in relation to particular target groups. Survey respondents identified support needs in relation to data on smaller under-represented groups such as care leavers and military families. They also outlined that it would be useful if the OfS was able to share examples of best practice to help higher-education providers deal with challenges associated with targeting at different stages of the student lifecycle.

- Training on the ‘nitty-gritty’ and the different techniques that can be applied in context. Much of the current guidance, while useful, is considered too general to support practical implementation of data and techniques on the ground. In-depth or tailored support is considered necessary to move some providers forward (whereby their ideas could be explored and refined in order to put in place a holistic approach to using data for targeting, monitoring and evaluation in context).

The development of toolkits for evaluation for practitioners that take account of their levels of capacity and expertise also emerges as a potentially useful approach to addressing data skills gaps. The model provided by the financial support evaluation toolkit would appear to be a pragmatic approach, working within the staffing and expertise constraints in institutions, which could be adapted for wider use in relation to evaluation of the outcomes to student success activities (subject to the provision of the associated outcomes data via HESA).

05. Legislation and ethical considerations

This section considers recent changes to data protection legislation and the impact that these have on the data landscape. It describes current systems and ways of working in terms of processing and sharing data, and highlights where systems and practices could be improved.

The General Data Protection Regulations

The General Data Protection Regulations (GDPR) developed by the European Union became part of UK law in May 2018.³⁵ The new legislation broadly follows the principles in the Data Protection Act 1998, but strengthens them in a number of important areas including: accountability and compliance; standards for valid consent; statutory obligations (including in instances of a data security breach); transparency; and the rights of data subjects. Higher-education providers collect, store, share and analyse personal data, including sensitive data, on both staff and students for a range of purposes, including targeting, monitoring and evaluating A&P activities. As such, the change in the law has had implications across the sector, resulting in the development of new systems and processes as well as policies and procedures to ensure compliance and that providers' obligations to the data subjects are met.

Legal bases for processing

Under the GDPR, it is incumbent on the higher-education provider as the data controller to identify one or more lawful bases for processing data and to determine what data should be collected, how it should be stored, who should be allowed to access it and how it should be protected. Establishing the legal basis is important because it affects both data-collection processes and data sharing. A written Data Sharing Agreement must also be in place between a data controller and a data processor with whom the data will be shared.

There are six lawful bases for processing personal data. Table 12 summarises the different legal bases for collecting and processing this information and key considerations for higher-education providers in the context of access and participation.

³⁵ In light of Brexit, the UK government has announced a new data protection bill and indicated that it will be substantially similar to the GDPR.

Table 12: GDPR legal bases.

Basis	Description	Considerations
Consent	An individual gives clear consent for their personal data to be processed for a specific purpose	This basis enables students to exercise complete control over what data is collected, by whom and for what purpose. The drawbacks for providers include administrative burden and low response rates.
Contract	Processing is needed in order to enter into or perform a contract	It can be argued that higher-education providers are processing data in order to fulfil a contractual obligation with the regulator. A tracking system could be under contract to a provider (Data Controller). This will not apply to the data that higher-education providers process with respect to their own students.
Legal obligation	Data processing is required in order to comply with the law (not including contractual obligations).	While higher-education providers are contractually obliged to set targets for A&P and monitor and evaluate progress, this is not enforceable by law and therefore this is not an appropriate basis for processing personal data in this context.
Public task	Processing is necessary to perform a task in the public interest or for official functions.	Higher-education providers fall within the scope of organisations performing duties in the public interest. This basis enables them to meet their statutory obligation to evidence the effectiveness of A&P activities to funders, regulators and central government and assure students and taxpayers that resources (including fee income) are being spent appropriately. Students do not specifically agree to their data being processed for this purpose.
Legitimate interest	Data processing is necessary for a legitimate interest unless there is a good reason to protect the individual's personal data which overrides those legitimate interests.	The interests of the higher-education provider has to be balanced against the interests of the student. While it is certainly in higher-education providers' interest to process data for the purposes of access and participation, it is less clear whether processing is in the students' interest. As with public task, students do not specifically agree to their data being processed for this purpose.
Vital interest	The processing is necessary to protect someone's life.	This is not applicable in the context of access and participation.

Different data providers have different legal bases for processing and sharing data. From a higher-education provider's perspective, this has implications for how data can be accessed, processed and used for targeting, monitoring and evaluation purposes. Table 13 provides an overview of the main differences between providers in relation to use of individual-level data.

Table 13: Legal bases used by data providers.

	Collection	Sharing
DfE (pupils)	DfE has legal powers to collect pupil, child and workforce data that schools, local authorities and awarding bodies hold. ³⁶	The law allows DfE to share pupils' personal data with certain third parties, including: schools; local authorities; researchers; organisations who make products connected with promoting the education or wellbeing of children in England; other government departments and agencies. ³⁷
UCAS (applicants)	Article 6(1)(b) where processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract, and Article 89 for statistical and research purposes.	Article 6 (1) (e) where 'processing is necessary for the performance of a task carried out in the public interest'. UCAS will only share personal information for purposes not directly related to their application to higher education where applicants have actively confirmed that they are happy for UCAS to do so.
HESA (students)	Article 6(1)(e) where processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the Controller, and Article 89 for statistical and research purposes.	HESA shares with third parties where there is a legitimate interest for statistical and research purposes. Information is supplied under contracts, which require that individuals shall not be identified from the supplied information without consent. HESA student information may be linked to school and/or further education college information and supplied to researchers where necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the Data Controller (Article 6(1) (e)). Data sharing is also for purposes of the legitimate interests of HESA in disseminating higher-education information, or the legitimate interests of third parties in undertaking research in the field of higher education (Article 6(1) (f)).

The three tracking systems also use different lawful bases for processing segmented data under the GDPR (Table 14). The decision on which basis is used may be informed by the requirements of data providers, such as HESA, to ensure tracking data can be matched to administrative data at the individual level.

³⁶ Section 114 of the Education Act 2005; Section 537A of the Education Act 1996; Section 83 of the Children Act 1989

³⁷ The Education (Individual Pupil Information) (Prescribed Persons) (England) Regulations 2009 (<http://www.legislation.gov.uk/ukxi/2009/1563/introduction/made>)

Table 14: Legal bases used by the tracking services

Service	Legal grounds referred to and roles
EMWPREP	<p>Basis for processing: Personal data, activity data and education data (including NPD/HESA matched data) will be processed under Article 6(1) (e) (public task) and Article 9(2) (j) (research purposes for special category data only). There is also a section regarding future context/involvement in additional research, which will be processed under Article 6(1) (a) (consent).</p> <p>Roles: Loughborough University is the data controller and members are data processors (and therefore required to use standard forms, privacy statements etc.)</p>
AWM	<p>Basis for processing: Article 6(1) (a) (consent) along with Article 6(1)(e) (public task) and research conditions</p> <p>Roles: the University of Birmingham, as the lead partner, is the data controller, but the other members of the Aimhigher Partnership also act as data controllers.</p>
HEAT	<p>Legal basis for processing: Article 6(1)(e) (public task)</p> <p>Roles: The University of Kent is both data controller and data processor, for aspects of HEAT activities. The individual data controllers (members) also specify how data will be used (lawful basis) and have a legal obligation to provide, or direct data subjects to, their own mandatory Data Protection Policy and Privacy Notices.</p>

The different requirements of the data providers also have implications for the legal bases that higher-education providers need to use in order to be able to match data at the individual level. This plays into the debate about the level of data used for monitoring and evaluation and the extent to which individual-level data, as opposed to aggregated information, is needed in order to draw robust conclusions (see Chapter 2).

Implications of GDPR

Stakeholders perceive that there has been a shift in students' and potential students' views since the introduction of the GDPR, influenced by broader public perceptions, both positive and negative, of how personal data is used by different organisations. Higher-education providers, therefore, recognise that there is an increasing need to ensure data subjects understand that their data will be used in ways that are both beneficial and ethical in the context of access and participation.

The introduction of the GDPR has benefited the data landscape in a number of ways. The increased focus on accountability and transparency has led to improvements in terms of:

- the level of clarity and agreement on what types of data need to be collected and why;
- the information available to data subjects, including on consent forms and in privacy notices that inform them of their rights; and
- the systems and processes in place for storing, processing and sharing data.

Ensuring compliance with the GDPR is not, however, without its challenges, and in some cases these are acting as barriers to effective data use. These are discussed in the sections below.

Protecting the interests, rights and freedoms of data subjects

In many instances, it has been necessary for providers to develop and implement new procedures to ensure compliance with the GDPR. This includes systems for communicating with and processes for responding to requests from data subjects, for example from those who wish to withdraw their consent, access the data held on them and/or 'be forgotten'. Developing these systems has involved an investment of time and resources and resulted in an increased burden on providers.

Interpreting the regulations

An underlying aspect of the data landscape for access and participation is the differing interpretations of the requirements at the local level, particularly when determining the legal basis for processing data. This has resulted in a range of approaches across the education sector which reflect the particular views of the experts and legal teams involved. Typically, interpretations of the regulations have erred on the side of caution. Many higher-education providers have chosen to process data on the basis of 'consent' as a result, in some cases at the insistence of wider stakeholders, such as schools and colleges (e.g. many schools and colleges insisted on learners giving informed consent to their data being processed for the purposes of the national evaluation of NCOP). While this has ensured the highest level of protection and control for data subjects, it has created uncertainty in terms of how data is processed for targeting, monitoring and evaluation purposes across the sector. It has created issues and challenges for higher-education providers and tracking services alike.

Consent

The use of consent as the legal basis is considered to be more resource-intensive than using other legal bases as it involves the implementation of administrative processes and the creation of an audit trail. Furthermore, in the context of access and participation, implementation usually relies on the co-operation of schools and colleges that act as intermediaries and obtain consent from pupils and parents on behalf of providers. Securing their co-operation, and supporting the implementation of these procedures, has time and resource implications for the schools and colleges themselves, as well as higher-education providers. The time required to obtain consent can impact on delivery schedules and, in some instances, deter schools and colleges from engaging in A&P activities altogether.

It can also be challenging to provide the depth of information required to enable individuals to give their full and informed consent in a succinct way, and in a way that students, particularly younger learners, can understand. There is a greater risk that consent is not obtained in this context, particularly from hard-to-reach groups (e.g. parents), or that the information collected is not complete. One stakeholder we consulted said that the rate of consent was as low as 13% at their institution.

Tracking learners across the lifecycle

Bodies such as the OfS can require or encourage higher-education providers to engage with data providers and intermediary bodies that have access to data that can be used for the purposes of targeting, monitoring and evaluating access and participation. For example, HEFCE set out an expectation that partnerships contracted to deliver the NCOP should engage with a tracking service in order to enable them to track learner participation in the programme, with a view to building a picture of successful interventions and approaches across the country.³⁸

Tracking organisations collect a large volume of data on learners who engage in A&P activities, which can be used to inform the monitoring and evaluation of outreach. However, a requirement to obtain consent to track learners and link activity data with other administrative sources can act as a barrier, as a lack of consent means services are not permitted to use the data they hold to track learners individually across the student lifecycle. A low consent rate is, therefore, considered to be a major threat to longitudinal tracking. For this reason, some higher-education providers have provided a ‘clear steer’ to tracking organisations that they should avoid using consent as the legal basis, if possible. In response, EMWPREP has developed an approach which will involve two legal bases for processing data going forward.

A consistent and well-coordinated strategy, supported by effective communication mechanisms, is perceived to help maximise the consent rate when consent is required. For example, EMWPREP and AWM have developed standardised consent forms for use by all members. AWM has also developed a standardised privacy notice and a partnership-wide data-sharing agreement, in addition to data-sharing agreements with schools and colleges. In addition to enhancing rates of consent, this centralised approach is perceived to be more cost effective and efficient to deliver.

The stakeholder consultations also highlighted that the tracking services have each come to arrangements with HESA whereby, for longitudinal tracking purposes, matched datasets can be created that allow for different levels of analysis of outcomes, depending upon whether the data subject (or parent) has consented. Where consent is in place, individual-level matched data is created by linking the tracking ID with the individual record; where consent is not in place, aggregate level or anonymised data is shared. This is particularly useful in the context of longitudinal analysis, whereby consent to track was not obtained prior to the introduction of the GDPR and it is not feasible or practicable to re-contact data subjects to obtain consent retrospectively.

³⁸ Source: HEFCE
<https://webarchive.nationalarchives.gov.uk/20180103171649/http://www.hefce.ac.uk/sas/ncop/resources/>

Discussion and future models

Stakeholders perceive a role for the OfS in offering some direction to providers on the most appropriate **legal gateway** for processing personal and sensitive data for the purposes of access and participation which avoids the need for participant consent. According to stakeholders, a stronger steer would reduce uncertainty and result in a more consistent approach across the sector. This would help to reduce the burden as well as the time and costs involved from the provider perspective; it would also facilitate opportunities for sector-wide or comparable approaches to using data in the targeting, monitoring and evaluation of A&P activities.

There is a need to develop effective **communications** which ensure data subjects understand the purpose for which their data is being collected and are able to make a fully-informed decision about whether to consent. These communications should reassure data subjects that data collection is appropriate, proportionate and ethical and that their interests, rights and freedoms are taken into account.

There is scope for expediting access to **data from schools and colleges, or local authorities**. This could involve, for example, clarifying that higher-education providers are within the scope of data-sharing agreements with partners such as educational establishments and local authorities, or facilitating access to data from the NPD.

The OfS also has an important role to play in clarifying (and strengthening) expectations in terms of monitoring. Clearly defining outputs and how they should be measured helps to minimise data requirements, which reduces the risk to all stakeholders and, more importantly, to students. This could also facilitate the development of more **centralised reporting structures** and ‘standardised’ data or ‘dashboards’ could be generated through the existing data collection and tracking systems.

06. Alignment of schools and higher education

This section considers current and potential future approaches to thinking about alignment between school systems and higher-education sector approaches to using data for targeting, monitoring and evaluation purposes.

Introduction

In addition to the main objectives of the review, the research set out to explore the potential for further alignment between the approaches adopted by schools and higher-education providers to using data for effective targeting, monitoring and evaluation. In the outreach space in particular, co-operation between higher-education providers and teachers is essential to the success of targeting and monitoring. Teachers are often the first point of contact and the gatekeepers to data and intelligence on potential participants, and are often directly responsible for making decisions on who participates, as well as maintaining an overview of participants' experiences and progress as a result of their engagement. The current research highlighted opportunities for closer alignment between school and higher education in two main areas: (i) the application of indicators and measures; and (ii) mechanisms for monitoring and reporting on success.

Indicators and measures

The existing literature highlights issues in relation to the extent to which the measures and indicators commonly used for A&P targeting are understood in schools and colleges. For example, recent findings suggest that POLAR (a higher education sector-related indicator of educational disadvantage) is not well understood by teachers, pupils and parents.³⁹ Interviewees in the current research did not highlight POLAR as a particular issue, but this was not a direct line of enquiry. There are also general misperceptions about whom outreach is for, running the risk that learners who are already on track for higher education are selected. Effective communication between outreach providers and teachers is vitally important to minimise deadweight issues. A recent report on provider perspectives concludes: "Being able to 'speak the

³⁹ Moore, J., Mountford-Zimdars, A. and Higham, L. (2018): Research into use of contextual data in admissions: Final report to the Fair Education Alliance (FEA), University of Exeter: Exeter.

language’ of schools and avoid prescriptive selection requests based on POLAR Quintile was seen as imperative to long-term engagement”.⁴⁰

Targeting is usually an iterative process involving negotiation and agreement. In all cases, the organisers of the activities will be aiming to establish a shared understanding with teachers of the target group(s) (while also taking account of the nature of the activities). At this level, the various stakeholders will decide which learners from disadvantaged backgrounds should be involved in higher-education outreach. Furthermore, the ultimate ambition could be that the outreach provision includes negotiation with target learners and their parents/carers (to enable them to be engaged in shaping their own educational future) – especially where providers are seeking to put in place a progressive programme of activity (i.e. working with targeted learners over a period of time). For the most intensive activities, there may be a focus on getting individual-level data via other sources than administrative data (e.g. using an application process). However, teachers often remain a vital link in this process. This can be affected by the quality of the relationships underpinning the collaborative arrangements in place. Teachers are also important gatekeepers because they will be able to help collaborative partnerships avoid multiple requests to the same learners and parents for personal information.

The school-level census (managed by the DfE) provides rich data on pupils in schools, including background factors and policy designations (e.g. FSM), plus educational progress and outcome measures. This data, along with ‘soft’ information on suitability and other background factors that might influence progression to higher education, is also held in schools. Although there is likely to be considerable overlap between some of the categories used, they are not common to all educational phases and schools and higher education may use different indicators to identify pupils who could benefit from additional support, including outreach, as a result. For example, there is a clear disjoint between a system which prioritises past geographical patterns of progression (POLAR), and a system based on levels of household income (defined in terms of eligibility for FSM).

To address this, tracking services have developed targeting models that align with school priorities, including indicators such as FSM/Pupil Premium. For example, EMWPREP draws on the locally-held data in order to develop lists of priority pupils who might be targeted for inclusion in the outreach by matching pupils to A&P criteria. The EMWPREP system includes a schools portal, which gives schools access to a postcode checker which can be used to assess postcode-related eligibility criteria (this is intended to be used by schools for which cohort lists cannot be sourced through their local authorities). AWM has a similar system in place, which sources learner data from schools and FE colleges to identify target cohorts. This data is subsequently made available to both outreach practitioners and schools/FE college

⁴⁰ Atherton, G., Boffey, R., and Kazim, T. (2019), How the targeting of learners for widening access to higher-education work could be improved. AccessHE: London.

staff. This process is supported by the central team to standardise the identification of target cohorts and monitoring engagement within outreach activities.

There are also more subtle disjoints due to a lack of shared definitions, as can be seen in the case of individuals with disabilities. For example, many more individuals have multiple disabilities than tend to be recorded/disclosed in higher-education data. Official statistics such as those provided by schools in NPD tend to be more reliable, although the SEN categories are based on pupils' greatest needs at school (which might not reflect various complex needs). Data collected by UCAS does not tend to facilitate distinctions between the severity of different types of learning challenges, and it is obviously hard in any case to compare people with different types of disability. Some authors have highlighted a further issue, which is that identification of learning difficulties may be linked to different approaches to diagnosing needs. Gorard & See (2013) identify that students in disadvantaged schools are more likely to be diagnosed with behavioural problems than those from more advantaged school contexts,⁴¹ although diagnosis of dyslexia has been on a rising trend.⁴² Another example is ethnicity, where there are different approaches to categorizing groups (HESA codes differ from DfE codes, for example).

The review suggests that providers see Free School Meal claims by Pupil Premium eligibility data as a particularly useful indicator for targeting. However, not all learners from low socio-economic backgrounds receive free school meals, plus there are issues around differences between those eligible for FSM and those receiving them. Therefore, while FSM may be useful as part of a basket of measures, it is unlikely to be a 'silver bullet', and should be reserved for targeting at the school rather than at the individual level.

The HEAT groups and composite model developed by AWM provide a designation that can be used on a practical level (as well as for reporting). This kind of approach, which takes in the level of disadvantage and prior attainment, is relatively easy to grasp and resonates with practitioners' understanding of under-represented groups. Anecdotal feedback suggests that head teachers can apply it to their thinking, and is especially helpful at distinguishing pupils (e.g. HEAT Group 2a (High Disadvantage, Highest Attainment) and HEAT Group 2b (High Disadvantage, Medium-High Attainment)). The approach is feeding into the way outreach is 'marketed' to schools to help them to identify participants, and this appears to be helping with a nuanced, shared conception of the target groups for outreach. Support for a measure of disadvantage contextualised by attainment was also seen as advantageous from a school's perspective.

⁴¹ Gorard, S. and See, BH. (2013) *Overcoming disadvantage in education*, London: Routledge.

⁴² Tomlinson, S. (2012) 'The irresistible rise of the SEN industry', *Oxford Review of Education*, 38(3): 267–286. <https://doi.org/10.1080/03054985.2012.692055>

Opportunities for shared monitoring and reporting

The push in schools towards personalised learning and teaching means that schools need to take a structured and responsive approach to pupils' needs, including making sure they are able to achieve and progress into higher education. There is an opportunity to reconcile school and higher-education data because of synergies between outreach delivered by higher-education providers, the school curriculum and schools' statutory duty to provide support on careers and progression. For example, it is possible to see links between higher-education outreach and the Gatsby Careers Benchmarks, which are being adopted by many schools as part of the government's careers strategy. There has been a move over a number of years to encourage schools to draw on the support of higher-education providers to raise pupil attainment and expectations of progression in education, including incorporating higher-education outreach into school improvement plans. Gatsby Benchmark dimension 7 directly relates to supporting pupil participation in outreach (Encounters with further and higher education), and there are synergies with other dimensions such as Learning from career and labour market information (dimension 2); Addressing the needs of each pupil (dimension 3); and Personal guidance (dimension 8).

There is some linking already going on between tracking services and schools. The system also allows teachers to view the activities their school has taken part in and produce their own report. Both EMWPREP and AWM are making links between school and higher-education priorities. A recent development for EMWPREP has been to incorporate the Gatsby Benchmarks into their system through its schools portal. There are plans to work with some pilot schools to explore how teachers might use the systems and what they would want the system to report. This information will potentially be useful to support school reporting to governors, Ofsted and others, including parents, on careers-related work.

It is important for both schools and higher-education providers to understand the policy context in which each is working. A current barrier to more joined-up approaches is different reporting and accountability requirements that put attention on different measures and criteria. More integration would put schools and higher-education providers 'on the same page' in terms of what is important from an accountability perspective. The proposal for new measures to be included in the school performance tables would seem to be an important opportunity for this going forward:

1. A schools accountability measure – a single score measuring progress from age 16 into higher education at age 18. It will compare previous attainment at GCSE and subjects studied against the national average for schools with similar student profile (due for publication in October 2019).
2. Longer-term destinations measure – will measure destinations after key stage 4 at three stages – after 1, 3, and 5 years (this measure is in development).

There is the opportunity for the development of a feedback loop which would emphasise the importance of collecting data from participants in A&P activities in order to check the effectiveness of the targeting processes at the school or college level. The monitoring data should ideally be used to inform the discussions between outreach practitioners and teachers (especially if the monitoring data suggests there is scope to improve the effectiveness of the processes in place to reach priority groups). This signals the need for an iterative process, which might slow the process down but should help to ensure that targeting works more effectively. This approach could also then be applied to tracking outcomes for evaluation purposes since schools hold a lot of outcome-related data – attainment measures being a main one – and in relation to higher-education applicants and acceptances, in cases where they are working with post-16 learners.

Discussion and future models

The areas of opportunity outlined above are inter-related in the sense that, taken together, they would constitute a mutually beneficial, system-level approach for targeting, monitoring and evaluation of outreach in the future. The introduction of a new measure of higher-education progression into the DfE performance tables seems to offer an ideal opportunity for higher-education providers to engage in more dialogue with schools, as part of a sector-wide conversation focused on agreeing shared approaches and new indicators for targeting individual pupils for participation in outreach (especially for the most intensive and sustained programmes of outreach).

Schools will play a pivotal role in any future model for outreach, and conceivably could take on more responsibility for both the application of targeting data and the capture of data for monitoring. Moving to a model where schools take on more responsibility for the application of criteria and ownership of outreach data would rely on high levels of school engagement, which would probably require some top-down directive, rather than relying on individual providers to lead. Data-sharing agreements would be needed, and there would need to be special arrangements in place where schools were in relationships with several higher-education providers.

Legal issues (as discussed in section 4) aside, data sharing tends to rely on the development of effective relationships with schools/colleges. Trust and shared vision for interventions are required, not only for data to be shared and used effectively, but also so that targeted participants are actively engaged and encouraged (rather than feeling stigmatized). The creation of a new policy designation for higher-education outreach could be a major step forward in creating the conditions where appropriate target groups are flagged up to practitioners on the ground. This could involve the application of new measures/composite measures, or a process of applying existing criteria to the individual pupil-level data. Taking action on the indicators and measures used to conceptualise the groups that are prioritised for outreach is probably the most immediate short-term aspect that the OfS could consider, especially as this links to the wider challenges of reliability in relation to the existing

datasets, as discussed. However, this will require further consideration to be given to the most appropriate indicators and datasets in collaboration with teachers and higher-education providers.

An integrated model of targeting, monitoring and reporting at school level would need the DfE to play the pivotal role in the application of data for targeting or, at least, opening up the data so that the individual target learners can be identified in the data. There would be an opportunity to do this consistently across the sector, benefiting providers and ensuring more consistent monitoring of school outreach.

Using tracking organisations or similar systems at the interface between schools and higher education could help to get buy-in from schools and promote shared monitoring/reporting functions. This would probably require standard data-sharing arrangements to be applied across all the higher-education providers and schools. The approach could allow for rich school-level data and information, including data on learner outcomes, to be captured.

07. Summary and recommendations

This research was designed to provide an initial mapping and review of the current data landscape, identifying how data is currently used for targeting, monitoring and evaluation of A&P activities. The barriers and solutions to optimising data use across the student lifecycle were also explored, with a particular focus on the role, and value for money, of the tracking services. This study is not exhaustive and some of the issues identified in the report require further exploration before definitive recommendations on future models and approaches can be made. However, we have identified key issues for consideration by the OfS and some initial recommendations for improvements to existing data systems and infrastructure. Brief guidance and recommendations for practitioners based on the findings has also been produced and will be published separately.

The recommendations outlined below are designed to inform the development and use of data to promote equal opportunities in higher education and to address the challenges, weaknesses and limitations associated with the current arrangements. Any new system that is developed should seek to achieve the following objectives:

- Proxy data and measures are as reliable as possible and ‘fit for purpose’ in relation to the objectives, context and stage in the student lifecycle they are used;
- Approaches to data collection and use across the sector are as consistent as possible;
- Duplication of data collection is reduced through data sharing (facilitated by common data definitions);
- Barriers to accessing linked datasets for tracking participant outcomes are minimised;
- The infrastructure for capturing the outcomes over the student journey is developed and maintained to ensure it remains fit for purpose in the context of a dynamic policy environment; and
- Institutional cultures that enable effective data governance, underpinned by appropriate institutional structures and processes, are encouraged and supported.

Short-, medium- and long-term recommendations are proposed here, structured according to the phase and stage of the learner journey. In this context, short-term is defined as within the next 6-9 months; medium-term, within the next 18 months, and long-term within the next 2 to 3 years.

Higher education providers' engagement with data

Targeting

Outreach

Datasets used for targeting need to be contextualised to local and institutional circumstances and, as such, a wide range of approaches to applying and using data has emerged across the sector. However, manipulating data into a format whereby it can be used to prioritise and target interventions at the provider level results in duplication of effort and inefficient use of resources. Intermediaries including the tracking services help to reduce these inefficiencies through, for instance, the provision of planning datasets. There is scope for greater sharing of effective methodologies for using data to target at different levels (school, individual). Data and resources, such as planning datasets, are not currently readily available in relation to all priority groups, including mature and part-time learners. There is particular scope to develop and share approaches to identify these, and other priority groups that are not represented in 'standard' datasets, such as white, working-class males.

The proxy indicators of disadvantage used to identify potential participants in A&P activities do not provide comprehensive coverage and vary in terms of their reliability. There is scope for the inappropriate application of the measures and/or 'cherry-picking'. Arguably, some groups who could benefit from outreach are not being engaged because they 'fall through the cracks' when a single proxy measure is used to inform targeting. Lack of a shared definition of type(s) of disadvantage and agreement on proxy indicators hinders the development of sector-wide perspectives. Systems for assessing individual-level criteria/eligibility vary and there are issues of poor external transparency and lack of resonance with partners in schools and colleges (i.e. measures are not understood/accessible to teachers, students and other stakeholders, sometimes coupled with lack of shared understanding of definitions of merit and potential for higher education).

Success and progression

The introduction of the A&P dataset has afforded higher-education providers a rich source of data, as well as focused attention on the whole student lifecycle. Providers are concerned about the timing of this data, the regularity of updates, and the availability of support to interpret the gaps data and/or recognise where progress is being made. Some would like simpler data based on a narrower set of measures, because complex data systems present challenges in providers with more limited data capacity.

Recommendations

For the purpose of targeting, we recommend that the OfS:

- Disseminate guidance for providers on indicators and measures, the circumstances in which they can be applied, and the inferences that can be drawn on the basis of them [short term];
- Develop a series of detailed case studies illustrating effective practice in the use of the A&P dataset for targeting (monitoring and evaluating) activities at provider level [short term];
- Explore options for a mature learner planning dataset (possibly drawing on the HEFCE AHE measure) [short/medium term];
- Convene a working group of institutions with a focus on mature and part-time learners tasked with making recommendations for new measure(s) of mature learner disadvantage, in order to inform the development of the new planning dataset (as well as future monitoring and evaluation of A&P activities for mature learners and other priority groups that are not included in standard datasets [short/medium term];
- Develop and pilot a new resource for school-level targeting that includes the new DfE measure(s) and matched NPD data for schools, with the objective of developing new sector-wide lists for targeting (building on tracking service planning datasets) [short/medium term]; and
- Establish a working group to develop and test a methodology for using the pupil-level datasets to prioritise learners for inclusion in outreach, which draws on the expertise of the OfS's analytics.⁴³ This would help to promote closer involvement between schools, higher education providers and analytical experts to review the availability of 'live' data required to most effectively identify target learners [medium term].

Monitoring

Providers collect a range of data to enable them to monitor the delivery of A&P activities. This is important as it enables higher-education providers to assess their performance against their A&P delivery plans and towards institutional A&P objectives associated with key performance indicators. Some higher education providers are only just starting to adopt efficient and effective methods of data management and governance.

Outreach

Higher-education providers focus on national A&P indicators, especially POLAR, for monitoring purposes. There is potential for greater alignment between priorities for schools and higher-education providers, and more integration between school and

⁴³ Composite measures are likely to be required taking account of attainment alongside factors of educational disadvantage. Further exploration of the experiences of providers utilising existing composite measures such as UCAS MEM and HEAT groups would be desirable to inform this work, as well as consideration of best practices in 'triangulating' data.

higher-education monitoring and reporting. Scope to measure and compare across institutions, to compare interventions and establish good practice, is limited by the use of different measurements or definitions. Tracking organisations play an important role for monitoring activity, because of the ease of reporting against targets and the ability to benchmark.

Success and progression

A&P practitioners have a limited understanding of available data, and how it can be used to monitor progress towards success and progression targets. Few higher-education providers have a framework or mechanism in place to address this gap through information and training. Current systems and approaches to monitoring vary greatly, although learner analytics is providing a useful tool in a number of cases.

Recommendations

We recommend that the OfS:

- Clarify expectations in relation to data for monitoring outreach, including monitoring the effectiveness of targeting, based on an understanding of the relevant consent and permissions required. This should include location markers in order to enable identification of ‘hot/cold spots’ [short term];
- Consider ways to minimise the cost to providers, particularly small and specialist providers, of accessing linked tracking data [short/medium term];
- Provide GDPR guidance on ‘best practice’ data collection for higher-education providers [short/medium term];
- Work with stakeholders, including tracking services, on the development of consistent definitions of A&P activities, standard reports and dashboards, potentially as part of its impact-reporting process for APP monitoring [medium term];
- Seek to minimise the barriers to accessing linked datasets for monitoring participant characteristics by expediting access to linked data and establishing the data-sharing basis for administrative data at the individual level to complement data that is captured through the trackers [medium term];
- Explore options for closer integration of schools and higher-education datasets for monitoring and evaluation, including incentives for schools to use tracking services (e.g. reporting against the Gatsby benchmarks) [medium/long term]; and
- Work with DfE to explore the feasibility of developing a single student identifier, such as the UPN, which remains with the student across the lifecycle to support data linking between phases [long term].

Evaluation

There is currently no system-wide approach to evaluating the outcomes/success of outreach in terms of increasing access to higher education. Evaluation usually requires understanding of the key outcomes of outreach interventions in terms of

application and participation rates to individual or multiple providers or higher education in general.

Outreach

Methodologies for assessing outcomes are being developed, including using different data sets such as applicant, student and linked datasets. UCAS applicant data is a common data source used by higher-education providers for evaluation purposes. Paid-for services such as UCAS STROBE can offer timely data but cost is a potential barrier. Tracking services have facilitated a review of where students progress to, regardless of which institution they received outreach activities from. Access to individual-level data on outcomes is hindered by different interpretations of GDPR requirements and the current lack of clarity on the legal basis for data sharing.

Success and progression

Current evaluation of success and progression activities is under-developed. Further work is required to strengthen evaluation practice and optimise the use of secondary data as part of a strategically-planned approach to demonstrate impact and understand the extent to which outcomes can be attributed to A&P interventions. Higher-education providers need support to develop a coherent evaluation plan and to devise appropriate methodologies that make best use of existing internal and external data sources. One approach could be to develop a toolkit for evaluating the outcomes from student success activities based on the model developed to support higher-education providers to evaluate their financial support. HESA has a key role to play in evaluating progression as the designated data body. HESA data will become increasingly timely as Data Futures comes on stream.

Identifying appropriate comparison groups against which to measure success is a key challenge for determining the effectiveness of interventions across the student lifecycle, due to issues of both access to data on comparators plus the diverse nature of the learners, contexts and activities involved.

Recommendations

We recommend that the OfS:

- Engage with tracking services to produce guidance on the use of tracking data for evaluation [short term];
- Work with DfE to explore how linked NPD data could be utilised through a data-linking service intermediary (such as a tracker organisation) to develop a standard methodology to facilitate comparisons with matched control/comparison groups [short/medium term];
- Work with providers and HESA to explore potential demand for, and the cost of, a new service for individual providers to evaluate whether interventions generate successful outcomes [medium term];
- Explore the feasibility of a pooled outreach activity dataset linking to matched data for tracking outcomes [medium/long term];

- Consider centralising the activities associated with data linking and reporting on the outcomes for participants in activities, either as part of an OfS function or via a third-party ‘Data Lab’ function, as exists in other policy areas [long term]; and
- Explore the potential for developing a toolkit to evaluate success activities based on the model developed for evaluating financial support, including the provision of the associated outcomes data via HESA [long term].

Higher-education provider engagement with tracking organisations

Tracking services offer a ‘bottom-up’ solution to the challenges of targeting, monitoring and evaluating outreach, by pooling resources and providing expertise and systems to support data use. Tracking service membership includes more than 50 providers and HEAT has the largest number of members. The system capabilities and types of support offered by tracking organisations are similar in many respects, but there are also a number of differences, with database functionality and reporting functions the two main distinguishing features.

Higher-education providers, irrespective of tracking service used, use the service to target and prioritise A&P outreach. However, none of the tracking services offer a complete solution to the effective use of data across the student lifecycle.

Tracking services facilitate access to linked data on individuals from the NPD. However, the tracking services do not support full coverage of A&P activity across the sector (the services do not include all activities of members and data from non-member provision is not included). As a result, it is not possible to accurately identify ‘cold spots’ within the current data landscape. This has implications for the effectiveness of targeting, as well as the achievement of A&P objectives at local and national levels.

Assurances that tracking systems are GDPR compliant is perceived to be a key advantage of using a tracking service. Access to training and support provided by tracking organisations is a further perceived benefit of membership to facilitate effective data use and analysis.

Overall cost, rather than perceptions of value for money, is the principal barrier to higher-education providers subscribing to a tracking service, which is likely to be affected by the capacity of the access and participation plan budget to absorb the financial burden; the extent to which the spending is perceived to be proportionate to the level of activity being delivered; and the existence of internal systems that are perceived to be sufficient for needs. Cost is, therefore, a prominent barrier for smaller and specialist providers. Our ‘value-for-money’ framework suggests the data infrastructure provided by tracking services is cost effective. Tracking services add value through a number of core features that cannot be easily replicated by individual higher-education providers. Access to training in data usage and analysis is also offered by tracking organisations.

Future models for tracking outcomes

Questions have been raised as to the value of having different tracking systems. Having multiple systems inevitably results in some duplication, but also leads to a

differentiated offer that meets the needs of different types of higher-education provider. Feedback suggests a preference for a range of options which reflect the diversity of the sector, rather than a 'one-size-fits-all' approach.

Recommendations

We recommend that tracking organisations:

- Work collaboratively and use consistent definitions of A&P activities [short term];
- Develop, together with the OfS, systems that provide clear communication to higher-education providers about the functionality and capability of tracking services [short term]; and
- Work collaboratively to develop more consistent standards of data collection [short/medium term].

We recommend that the OfS:

- Considers ways in which tracking membership could be accessed by a wider range of providers, in particular small, specialist providers [short term].

Legislation and ethical considerations

The introduction of the General Data Protection Regulations (GDPR) has had implications for the collection, use and storage of data by the higher-education sector. The new regulations have led to improvements in procedures for ensuring informed consent and for storing and processing data, and have increased clarity and agreement on what data needs to be collected and why. The changes have had resource implications and have resulted in an increased burden for higher-education providers, although this may not persist in the long term.

However, there is still a lack of clarity on the legal basis for data collection, and providers are looking for further clarity on GDPR requirements to ensure they comply with the new legal requirements for data protection and processing without increasing burden and adversely impacting data collection. Access to data can depend on relationships with schools and colleges. Reliance on consent as the legal basis increases the administrative burden and can lead to a reduction in the data available. The capacity to obtain consent from young people and/or their parents, and put in place systems to manage data processing and update consents, etc, over time, differs by provider. This can act as a barrier to the type of data and length of time that data can be retained across the student lifecycle, which can impede longitudinal tracking.

Recommendations

We recommend that the OfS:

- Work with the ICO to clarify the legal basis for data collection, processing and matching and the basis on which matched data can be shared with data providers, including HESA, for the purposes of evaluation [short term];

- Once clarity has been obtained, produce guidance on best practice for higher-education providers, drawing on the data protection toolkit developed by DfE [short/medium term]; and
- Explore, together with higher-education providers and HESA, potential demand for, and the cost of, a new service for individual providers to evaluate whether interventions generate successful outcomes [medium term].

Alignment of schools and higher education

Cooperation between schools, colleges and higher-education providers is essential for effective targeting, monitoring and evaluation of outreach interventions. The current research highlighted opportunities for closer alignment between school and higher-education providers regarding the application of indicators and measures and in relation to the mechanisms for monitoring and reporting on success. Effective communication between outreach providers and school personnel can help to minimise deadweight issues associated with ineffective targeting of which learners should receive particular interventions. Schools could conceivably take on more responsibility for both applying for and capturing data that can be used for targeting and monitoring purposes. An integrated model of targeting, monitoring and reporting at the school level would require the DfE to play a pivotal role including, as a minimum, opening up access to the data so that individual target learners can be identified. Using tracking organisations or similar systems at the interface between schools and higher education could help to secure buy-in from schools and promote shared monitoring/reporting functions.

Recommendations

We recommend that the OfS:

- Encourage higher-education providers to make A&P measures more accessible to stakeholders and students (including look-up tables where possible) [short term];
- Work with stakeholders to develop a methodology to support schools to target groups at risk of not fulfilling their higher-education potential [short/medium term]; and
- Encourage the sharing/matching up of identifiers between phases by promoting the wider use of the Unique Pupil Number (UPN) [long term].

Appendix 1: Methodological note

Desk research

Desk research was undertaken to enable the mapping of the current system in terms of infrastructure, stakeholders and data flows, while taking account of existing arrangements and protocols, future proposals (e.g. open data strategies) and supply and demand codes of conduct. Research, analysis, A&P consultation responses, sector guidance, funded programmes, and best practice materials were included in the review to provide insights into: the policy context; current arrangements for using administrative data for targeting, monitoring and evaluation; evidence of the strengths and weaknesses of current approaches (and implications for A&P provision). This was supplemented with a review of the strategies and plans of data providers and services. Of particular relevance were Data Futures protocols and open data strategies. The research objectives provided the parameters for the review, which focused on materials produced in the last 5 years to ensure it was sufficiently relevant to the current context. Given the short timeframe for the research, the material was selected in consultation with the OfS and the steering group.

A review of data protection legislation and associated guidance and ethical guidelines (in particular the introduction of GDPR) was carried out to clarify the requirements for accessing and sharing data. The implications for current and future systems and processes to support the use of administrative datasets for targeting, monitoring and evaluation were also considered.

Stakeholder consultation

A total of 32 semi-structured telephone interviews, on an individual or group basis, were carried out between April and July 2019 with a range of senior stakeholders. This included administrative data providers (11 interviews in five data provider organisations), data users (17 interviews across 11 A&P provider organisations including universities, colleges and a school) and tracking system providers (4 interviews).

Administrative data provider interviews

Representatives from UCAS, SLC, ImpactEd, the OfS, DfE (NPD, ILR, LEO), and HESA participated in the consultation. The key objective of these interviews was to capture and assess information about current arrangements for using administrative datasets for targeting, monitoring and evaluation; levels of engagement with higher-education providers, data flows and data constraints faced by providers; opportunities for alternative models to inform the future development of infrastructure; legal and organisational requirements; value for money / value added; and insights about the future data landscape for effective use of data to promote equal opportunities.

Data-user interviews

Seven higher-education providers, a college, TASO, the Sutton Trust and United Learning were consulted. Perceptions and experiences of using data for targeting, monitoring and evaluation of A&P activities, the costs and benefits involved, the benefits and impact of using data for their provider and the sector, and challenges and opportunities of effective data use to support A&P objectives were all explored. Consultations with data users were used to inform models of data use, capacities and emerging issues that were further explored via the data-user survey.

Drawing on the range of institutional experiences enabled an assessment of the perceived value of current services and whether changes to existing arrangements, or an alternative system, would deliver better value for money by ensuring more effective and efficient use of administrative datasets.

Tracking system provider interviews

Depth interviews with tracking system providers (AWM, EMWPREP, HEAT) explored perceptions about the current role and expectations of the tracking services; their perceived strengths and weaknesses, including evidence of their effectiveness; and the costs and value to individual A&P providers and the sector. The ‘added value’ of tracking system services offered was explored including data capture and management, data analysis and longitudinal tracking. The products and services developed for members, and the role they play in training/capacity building, sharing, innovation and effective practice were also explored.

Data-user survey

Findings from the desk research and stakeholder consultation were synthesised to inform the design of an online survey of data users to capture a broad spectrum of views from staff within higher-education providers involved in the use of data for targeting monitoring and evaluation A&P activities. Gaining perceptions of data users on a broad scale provided the opportunity to determine the general level of engagement with administrative datasets and the impact. The survey was designed to enable the project team to understand the range of different approaches used by data users to obtain and engage with administrative datasets for targeting, monitoring and evaluation purposes. The potential added value of using a tracking service was also considered, in addition to collating insights from non-members of tracking services.

Two online surveys were developed to explore perceptions and experiences of how organisations use data and information to inform their targeting, monitoring and evaluation of: (i) A&P **access** to higher education-related activities; and (ii) student **success and progression** activities.

Unique survey links were disseminated by the OfS on behalf of CFE between 18th July and 11th August 2019, to around 260 organisations listed on the OfS register of English Higher-Education Providers database. In the absence of a comprehensive sampling frame, a ‘snowball’ technique was used to disseminate the survey including

direct approaches, networks, mailing lists and social media. Unique survey links were emailed to a lead contact at each organisation, together with a ‘frequently asked questions’ document that outlined the project aims and objectives and which staff the survey links should be forwarded to. A total of 82 unique providers responded to at least one survey, representing a response rate of 45 per cent. A total of 76 responses to the ‘access’ survey and 60 to the success and progression survey were received.

Survey analysis

Top-line analysis was carried out to explore the range of data sources that higher-education providers use, perceptions of usefulness, and perceived barriers for A&P targeting, monitoring and evaluation activities. Tracking organisation membership (member / non-member) and provider size. Provider size was defined by A&P 2019/20 projected expenditure. The sample was divided into 3 equal groups; small providers under £690k (n = 23), medium-sized providers with median level expenditure of up to £7.1 million (n = 24), and large providers with the top third expenditure of up to 18.7 million (n=25). Cross-tabulations were undertaken using these variables in order to explore differences in perceptions and experience of data use in practice.

Appendix 2: Examples from other sectors

Name of Service	Ministry of Justice Data Lab
Link	https://www.gov.uk/government/publications/justice-data-lab
Aims	Improve research and evaluation capability for organisations delivering offender services by allowing access to high quality re-offending data.
Who uses it	Organisations who have worked with offenders (the users include charities; public sector organisations; private sector organisations; educational institutions).
Processes applied?	The organisation completes a template that asks for details of the individuals and activities delivered (needs to be at least 60 individuals). The data is sent using a secure email account (e.g. CJSJ, GSI, and PNN). The data is matched to MoJ (the 'treatment group'). A matched control group is constructed using Propensity Score Matching. Re-offending is compared between treatment and control group. Analytical results are reported back to providers using a Report Template. The matched control group aims to control for other factors that might influence the likelihood of an offender receiving the intervention and/or re-offending (e.g. criminal history). The report template includes a statistical significance value. Reports are published as official statistics.
Underpinning datasets?	Police National Computer (PNC) managed by the Home Office, reoffending data, Department for Work and Pensions (DWP) / HM Revenue and Customs (HMRC) data, Offender Assessment information. The Ministry of Justice receive monthly extracts of data from the PNC. The data has also been matched with administrative datasets from DWP and HMRC, to provide information about offenders' benefit and P45 employment history.
Fields/Indicators used?	<p>The one year proven re-offending rate (defined as the proportion of offenders in a cohort who commit an offence in a one year follow-up period which received a court conviction, caution, reprimand or warning during the one year follow-up or in a further six-month waiting period, starting when offenders leave custody or start their sentence in the community).</p> <p>Frequency of reoffending, time-to reoffending and total number of offences, and severity of reoffending is now included in the most recently published reports in terms of three court outcome levels, which classify offences in three levels based on legal criteria as indictable-only, triable-either-way and summary.</p> <p>Most recently they have also added outcomes related to employment, job retention and receipt of benefits.</p>
Individual level data?	No personal information is made available via the Justice Data Lab (JDL) service. Once individual level data is submitted to the JDL, it is linked to various datasets before aggregated outputs are provided.
Legal basis	The legal gateway is Section 14 of the Offender Management Act 2007 (which permits disclosure of information for the purposes of the management of offenders). MoJ relies on Article 1(e) (processing necessary for the performance of a task carried out in the public interest

	<p>or in the exercise of official authority vested in the controller). As data relating to criminal convictions and offences is being processed Article 10 of the GDPR and section 10 and schedule 1 paragraph 4 of the DPA 2018 apply.</p> <p>The users are the data controllers and need to give legal assurance that the transmission of data is compliant with data protection law and ensures that analysis produced through the Justice Data Lab is used accurately. The legal basis is either via obtaining consent directly from offenders or (most likely route) via Article 5 1 (b) (public interest/research).</p>
Resourcing	<p>The service is delivered by a small team in the Justice Statistics Analytical Services (JSAS) (initially two people increasing to four people in 2014).</p>

Name of Service	Data Access Request Service (DARS), NHS Digital (formerly HSCIC)
Link	https://digital.nhs.uk/services/data-access-request-service-dars https://digital.nhs.uk
Aims	The aim of NHS Digital is to harness the power of information and technology to improve health and care. The DARS service has been set up to support access to data by stakeholders to help improve NHS services.
Who uses it	Clinicians, researchers and commissioning organisations.
Processes applied?	Users go through an enquiry and application process. If successful, the Data Sharing Agreement is electronically signed by NHS Digital in DARS Online. The data, with patient objections upheld as appropriate, is produced, reviewed and signed-off by NHS Digital, or the data service access is granted. The data are made available either by secure file transfer or through the Data Access Environment.
Underpinning datasets?	Range of datasets including Hospital Episode Statistics (HES); Patient Reported Outcome Measures (PROMs); Office for National Statistics (ONS) Mortality Data; NHS Registration Data from the Personal Demographics Service (PDS) plus other data such as on Diagnostic Imaging, Mental Health, and data from surveys and measurement programmes.
Fields/Indicators used?	Depends on the dataset and includes for example patient outcome measures, plus records of patients attending accident and emergency units, admitted for treatment or attending outpatient clinics at NHS hospitals in England.
Individual level data?	Patient-level data can be made available to organisations that meet the Information Governance (IG) requirements. Identifiable data may be provided where appropriate. The service also provides bespoke data linkage (linking two or more datasets on DARS or linking customer data to data that held by DARS, or two or more sets of customer data). DARS can provide patient status and tracking (the demographic status of a specific group of patients or tracking them over a period of time, and providing regular updates using PDS, ONS cancer data and ONS mortality data). DARS can match individuals to provide: validation of demographic data to improve linkage outcomes and/or ensure it is accurate prior to patients being contacted; a snapshot of current demographic status and mortality including cause of death where appropriate; periodic long-term updates on demographic status, morbidity and mortality of a patient cohort.
Legal basis	As an executive non-departmental body reporting to the Department of Health and Social Care, most of the processing is directed by the Secretary of State. Legal basis for dissemination is the Health and Social Care Act 2012 - s261 (1). Under the terms of the data-sharing agreement, NHS Digital consents to the appointment by the Data Recipient of the specified party to act as its Data Processor solely for the processing activities set out in the agreement.
Resourcing	Charges apply. Indicative set up and first-year service charge is £1,000 and Annual Service Charge is £1,000. See

	https://digital.nhs.uk/services/data-access-request-service-dars/data-access-request-service-dars-charges-2018-19 for details of charges for products.
--	--

Name of Service	'Clinical Trials Service', NHS Digital (formerly HSCIC)
Link	https://digital.nhs.uk/services/clinical-trials-service https://digital.nhs.uk
Aims	NHS Digital is the national provider of information, data and IT systems in health and social care, with a remit for ensuring information flows efficiently and securely. NHS Digital and Health Data Research UK agreed an approach to using data to better inform clinical trials and to design systems that allow data from across the English healthcare system to be accessible for clinical trials.
Who uses it	Data is used in relation to trials with scope of Health Data Research UK, the national Institute for data science in health and associated research initiatives. A recent example is the use of data project on clinical trial recruitment led by University of Oxford Biomedical Research Centre.
Processes applied?	A suite of services is being developed to support clinical trials with the aim of speeding up the recruitment process and reducing the number of recruitment centres, cost and complexity of late phase trials, and achieve comprehensive data collection for participants. The services include feasibility testing (using national datasets to provide numbers of those eligible for a clinical trial based on certain criteria); Identification of potentially eligible participants (using clinical and demographic datasets for possible recruitment into a trial); Tracking for contact details (i.e. communication with all trial participants during and after the trial); Tracking for clinical outcomes (regarding major life events for trial participants).
Underpinning datasets?	As appropriate to the aims and objectives of the particular trial.
Fields/Indicators used?	As appropriate to the aims and objectives of the particular trial.
Individual level data?	As appropriate to the aims and objectives of the particular trial but could include identification of potentially eligible participants; tracking for contact details and clinical outcomes.
Legal basis	Section 254 of the Health and Social Care Act enables the Secretary of State for Health and Social Care to direct NHS Digital on matters concerning the provision of health services or adult social care in England. The Secretary of State can instruct NHS Digital to: Put in place systems to collect and analyse information; Develop or operate information or communication systems.
Resourcing	Not specified.

Appendix 3: Overview of the functionality of the tracking services

		Aimhigher West Midlands	EMWPREP	HEAT
Organisational aspects	Lead/Hosting arrangements	Led and hosted by University of Birmingham	Led and hosted by Loughborough University	Led and hosted by University of Kent with peripatetic staff hosted by member institutions
	Members	Six Higher Education Institution (HEI) and six Colleges of Higher Education (CHE). Location in the West Midlands and South West	37 members (ten full HEI partners; 3 half HEI partners; 4 NCOP Partnerships; and 20 CHE providers) Located across the Midlands	79 HEI members; 3 third sector member; 24 NCOP partnerships located across England
	Cost to members	The cost is part of a wider service arrangement, which includes access to the activities delivered by AWM (e.g. residential activities and mentoring programmes) as well as entitlement to use the data systems and support services (£35,0000 per partner). In terms of data services, the employment costs relating to the data services side of the operations are estimated to equate to £4,333pa per member HEI	'Full members' pay £10,500 pa (2-year contracts) NCOP partners are 'Half members' at £5,000 pa ⁴⁴ (FE colleges £2,000)	£6,450 pa for institutional and third sector members NCOP partnerships pay 0.65% of the respective NCOP partnership budget to use the services (in NCOP phase 2) ⁴⁵
	Other income	-	-	(HEFCE) Catalyst Funding totalling £3million from September 2014 to September 2017

⁴⁴ HEIs that were already part of NCOP did not need to pay a subscription to use the service. Those that weren't part of EMWPREP were asked to pay 'half' subscription to use the service. HEIs can top up that funding to full membership if they want to.

⁴⁵ The charge has increased in phase 2 (based on the patterns of demand in phase 1).

		Aimhigher West Midlands	EMWPREP	HEAT
	Staffing	One full-time member of staff is employed for database development and training/support across all types of members (the wider AHWM team comprises a manager, lead officer, evaluation assistant officer and finance/admin assistant)	Currently the Co-ordinator and Data Officer. When at full capacity the team comprises the two full-time and two 0.65 FTE staff	The central team comprises the Director, Service Manager, two Member Consultants, two Data Analysts, a Support Officer and Communications Officer. The team also draws on a business analyst and research consultant
	Governance/ Structures	Strategy group comprises all institutional partner representatives (PVC level), which oversees the work of a management group and various operational groups	The partnership steering group (with equal member voting rights) agrees the EMWPREP team's annual work programme	A steering group (comprising representatives from 25 members on 3-year terms with an independent chair) steers the development of the services and has budget oversight. The Governance Board has oversight of the overall delivery and model (the lead institution and five HEI representatives). All members are consulted via a formal consultation process for their views in relation to decisions on services
Targeting/monitoring datasets	School/college level datasets/Look-ups	School and college level profiles Analysis by MSOA using NPD (currently via HEAT)	School and college level profiles Analysis by MSOA using NPD (currently via HEAT)	School and college level profiles Analysis by MSOA using NPD
	Individual level datasets/Look-ups?	Postcode look-up checker (shared with EMWPREP). Work is planned in NCOP phase 2 to enable scoring of learners against self-reported aspiration or attainment markers in order to target specific interventions and monitor progress using a learner analytic based approach (the Learner Evaluation and Progression (LEAP) toolkit)	Post-code look-up checker (shared with AHWM) Work has been completed to create an analysis tool for a partner which scores students according to a range of criteria obtained through application information	Postcode profiler tool

		Aimhigher West Midlands	EMWPREP	HEAT
	Systems for local data sourcing?	For NCOP collaborative delivery, data is requested from schools in order to create activity level registers, and individuals are then associated with the activities on the system they take part in.	There is an annual programme of work with five local authorities to identify individual target learners in year 7-13, firstly, using socio-economic and 'potential' data (subject to data sharing agreements), and secondly according to NCOP criteria, and providing this information to schools and colleges to assist them in the WP targeting process	
	Other types of participant data held at individual level?	Baseline questionnaire results (e.g. aspirations, awareness of student finance) Activity related questionnaires (based on five As to progression to higher education, awareness, aspiration, attainment, application, and access) Post-activity questionnaires	Data from standard and institutional activity related questionnaire surveys	Data from standard and institutional activity related questionnaire surveys HEAT group category information Student profiler tool: allows members to see whether a student has participated in other member outreach (data restricted to minimal detail)
Arrangements for processing data	Roles	In relation to Aimhigher programmes, the University of Birmingham, as the lead partner, is the data controller, but the other members of the Aimhigher Partnership also act as data controllers	Loughborough University is the data controller and members are data processors (and therefore required to use standard forms, privacy statements, etc.)	The University of Kent is both Data Controller and Data Processor, for aspects of HEAT activities.
	Legal basis for processing	Article 6(1) (a) (consent).	Personal data, activity data and education data (including NPD/HESA matched data) will be processed under Article 6(1) (e) (public task) and Article 9(2) (j) (research purposes for special category data only). There is also a section regarding future context/involvement in additional research which will be processed under Article 6(1) (a) (consent).	Article 6(1)(e) Necessary for a task carried out in the public interest The individual Data Controllers (member) also specify how it will be used (lawful basis) and have a legal obligation to provide, or direct data subjects to, their own mandatory Data Protection Policy and Privacy Notice etc.

		Aimhigher West Midlands	EMWPREP	HEAT
	Retention of data	Until participant is 25 years of age	Full information is held for the lifetime of the partnership (currently until July 2021). Once matched to school and university records identifying details will be removed from the record. An example of the current duration length for HESA matching can be found here in our Data Retention and Deletion policy document.	The individual Data Controllers (members) need to specify how long data will be retained
Longitudinal tracking of learners	Administrative datasets used for tracking?	Monitoring participant outcomes in terms of their subsequent higher education participation using student data from HESA National Pupil Databases (current request is outstanding)	Monitoring participant outcomes in terms of their subsequent higher education participation using student data from HESA	Key stage 4 attainment by linking student records through the NPD Progression to Level 3 study through the NPD Monitoring participant outcomes in terms of their subsequent higher education participation, success and progression using student and graduate data from HESA. Includes success rates in higher education, PG study and employment destinations (DLHE data).
	Other mechanisms for tracking?	Use of school level UCAS information (i.e. via schools and colleges. This is included in the local collaborative agreements). Nb. The focus is on tracking NCOP learner participants (rather than institutional work per se)	Monitoring participant outcomes in terms of their subsequent Key Stage 4 attainment and progression to further education (subject to data sharing agreements with local authorities)	
	Approach to working with comparison groups	For NCOP work, comparison groups are NCOP AHWM learners who, for whatever reason, haven't engaged in the programme activities.	This has not been requested by members although may emerge as potential focus in future.	

		Aimhigher West Midlands	EMWPREP	HEAT
Training/consultancy services	Systems related support	Regular database training sessions (approximately six-monthly) Ad hoc informal support as requested	Training on the database as required by new members Ad hoc group training sessions Ad hoc informal support as requested	Initial set-up training (at the start of membership on use of the HEAT Service) User access to on-going training college with user guides, materials, videos and webinars ⁴⁶ Online training sessions Members may be charged for additional training that is outside the scope of set up and online training Helpdesk facility User forums/Data Users' Network
	Evaluation related support	AHWM has developed evaluation toolkits including pre- and post-activity questionnaires for NCOP work	Designing and implementing bespoke research and data methodologies as per partner requirements (using quantitative, qualitative or mixed-methods approaches). Provision of full specific intervention evaluation reports have included focus groups with young people taking part in outreach activities and interviews with teaching professionals and other key stakeholders (limited to one evaluation per HEI a year)	Research Group Integrated survey tool: to design and deliver electronic surveys for evaluation purposes (linked to participant and activity data) (Forthcoming) Evaluation tool: An online tool to assist with evaluation planning
	Other member communication mechanisms/infr astructure	Monthly operational updates (which include data related issues as appropriate)	Bi-monthly newsletter The Director of EMWPREP runs the NEON impact and evidence working group	Member forums and networks Integrated file store to allow access to shared documents and data across the

⁴⁶ Examples of sessions have included: Analysis of HEAT data - for members wishing to learn how to do their own bespoke analysis to supplement standard HEAT reports; Infographics – for those members wishing to learn how to create visuals to support the communication of their research outcomes.

		Aimhigher West Midlands	EMWPREP	HEAT
				membership and individual member files by secure access
Reporting	Reporting cycles	Tri-annual reports against KPIs to correspond with NCOP management group meetings	Formal reports collated twice a year in February and August and disseminated in March and September	Annual reporting (linked to timing of access to outcomes through matched datasets)
	Member levels reports	Periodic reporting against partnership KPIs Reporting tools for members are part of the database	Preparation of interim (in-year) and annual HEI level reports, and associated infographics Preparation of interim (in-year) and annual NCOP level reports (Forthcoming) Plans exist to put infographics into the database to enable members to access reports	Three levels of annual reporting: Report 1: Level 2: KS4 attainment of students taking part in outreach activity recorded on the HEAT database Report 2: Level 3: Progression to Level 3 study. Report 3: Level 3+: Progression to higher education study. Includes success rates in higher education, PG study and employment destinations. Members access/download own reports via the database/dashboards
	Membership wide reporting			Research reports using Aggregate Results of all member data (as context for participating institutions). Intersectional analysis of outreach participant outcomes (e.g. HEAT group reports)
Lifecycle related	Services relating to student success		Provision of contextual data via the postcode database Analysing and reporting partner institutions' student data: widening participation trends of HEIs' applications and admissions data; comparing	By linking participant data to student outcome data HEAT can explore in detail the relationship between outreach participation and student success across the whole student lifecycle

		Aimhigher West Midlands	EMWPREP	HEAT
			offer/acceptance rates of WP/non-WP students; monitoring and reporting student retention trends of WP/non-WP students	
	Student progression-related services			Matched progression data sourced through HESA includes success rates in higher education, PG study and employment destinations
Other	Other services/ functionality offered to members			Registration tool (online tool to facilitate capture of registers/participation details electronically at point of delivery) Student Ambassador Portal and System (SAP) (online system integrated with HEAT database to facilitate recruitment and management of Student Ambassadors) Online Events Programme (OEP) (electronic invite, acceptance and capture of data from teachers and learners)

Appendix 4: Barriers to effective targeting, monitoring and evaluation

Barrier	Targeting Access (base = 76)	Targeting S&P (base = 60)	Monitoring Access (base = 76)	Monitoring Success (base = 60)	Monitoring Progression (base = 60)	Evaluation Access (base = 76)	Evaluating Success (base = 60)	Evaluating Progress (base = 60)
Gaining access to appropriate datasets/systems	66	33	\	\	\	58	43	47
Getting data from participants	71	\	\	\	\	\	\	\
Disclosure of information issues	\	40	\	\	\	\	\	\
Accessing required data from within a dataset	38	37		\	\	38	30	38
Lack of staff time	57	57	51	38	38	66	43	42
Lack of expertise in data analysis and interpretation		35		\		46	35	33
Lack of accurate data	\	22	\	\	\	29	\	\
Lack of benchmarked data	\	43	\	\	\	55	\	\

Barrier	Targeting Access (base = 76)	Targeting S&P (base = 60)	Monitoring Access (base = 76)	Monitoring Success (base = 60)	Monitoring Progression (base = 60)	Evaluation Access (base = 76)	Evaluating Success (base = 60)	Evaluating Progress (base = 60)
Lack of expertise in using the system	\	\	37	20	25	\	\	\
Poor alignment of system with your organisational priorities	\	\	29	22	20	29	20	17
Poor data coverage	45	\	29	30	28	38	25	37
Lack of knowledge of data sources/ systems that can be used	36	30		\		43	45	42
Lack of knowledge of systems that can be used to support monitoring	\	\	39	33	37	\	\	\
Cost of data	37	22	53	32	33	49	27	28