Widening the evidence base: use of broader data and applied analytics in NICE's work

Scope

This document sets out:

- NICE's ambition to use broader sources of data and analytic methods to enhance our existing methods and processes **and**

- how we intend to do this, so that stakeholders are informed and can engage with us when we implement these proposals.

We have not included a detailed description of methodological aspects and process considerations here, as these issues will be covered in detail in future papers.

Terminology

In this document the term 'data' refers to any source of quantitative or qualitative data that is suitable for use in NICE's work programmes, when examined using a range of analytic techniques. This definition includes (but is not limited to):

- electronic health record data

- 'real-world data' looking at health and social care practice outside of trials (for example, registries)

- any other relevant data that has been made available for others to use.

The term 'data' does not, in this context, refer to published research findings and summary statistics. These will continue to form the core evidence for NICE recommendations and advice, and this paper sets out ways that other types of data may be used alongside these sources.

Contents

Introduction

1. NICE helps the health and social care system to deliver the best outcomes for people using services with the resources available. We do this by developing recommendations, advice and information through a diverse range of programmes, which share the same core process of identification, assessment and interpretation of evidence.

2. The recommendations we make and the information we provide all need to be kept up to date, requiring periodic repeats of the guidance development process, or variations of it.

3. NICE's ambition is to embrace a range of technological advances and analyse a broader variety of data to develop evidence to inform our work. Fundamentally, the suitability of any particular data as evidence will be determined by the quality of the data and the nature of the question that we are trying to answer.

4. There are 5 key strategic areas that underpin the use of data analytics across NICE: data, tools, skills, collaboration and public trust. With an increasing emphasis on using data analytics across all our guidance, we will potentially be able to:

    -   generate new evidence through data analytics, to help develop and update guidance more efficiently than is currently possible

    -   answer questions that we cannot answer using our traditional approaches (for example, when systematic reviews identify gaps in the current literature, when considering questions for which trials are unlikely to be carried out, or when existing trials do not include populations of interest)

    -   monitor and validate intermediate outcomes

    -   measure the effectiveness and cost effectiveness of interventions in real-world settings

    -   improve our tracking of guidance implementation, uptake and impact, and our use of this information to inform the need to update guidance.

5. This will be achieved through:

    -   unlocking and exploiting the full potential of data from a range of sources, including health and social care organisations, electronic health and care records, research cohorts and patient surveys

    -   developing partnerships with a range of expert organisations across the health and social care system and the wider data community

   -   increasing capability to link live systems and unstructured data

- translating data into evidence, recommendations and innovative interactive tools, which allow health and social care professionals and the public to put our analysis at the heart of decision-making.

6. There is some overlap between this document and the [Centre for Health Technology Evaluation methods update](#), in particular around areas of uncertainty, types of evidence, and evidence generation. However, this document also covers work in other NICE programmes (such guideline development). We will aim to keep our processes and methods consistent across NICE when possible, but they may vary depending on the contexts in which data are being used.

The role of NICE

7. Since 1999, NICE has provided guidance and advice to the NHS. NICE was initially set up to reduce variation in the availability and quality of NHS treatments and care. However, over time our remit has expanded to include producing guidance to prevent ill health and promote healthier lifestyles, and developing guidance and advice for social care.

8. NICE guidance is officially for England only, but certain products and services are provided by agreement to Wales, Scotland and Northern Ireland.

9. NICE develops a range of different types of guidance and advice. This includes:

- guidance for the NHS on health technologies (including procedures, diagnostic agents, devices and pharmaceutical and biopharmaceutical treatments) through our Health Technology Evaluation programmes, which ensures that people can access available treatments that are clinically and cost effective

- advisory guidelines for health and social care professionals, which demonstrate best practice for diagnosing and managing a range of conditions, supporting people's social care needs, and encouraging better health

- advice products that evaluate published evidence on a range of topics

- standards and indicators to support delivery of high-quality care.

Key NICE guidance, standards and indicators are illustrated in figure 1.

Figure 1: NICE guidance, standards, and indicators. *NICE guidance consists of evidence-based recommendations which are developed by independent committees. Quality Standards are used to identify areas for quality improvement in health and social care. The NICE Indicators menu is a set of measures based on outcomes or processes for health and care. The aim is to give everyone access to high quality care and provide best value for the NHS and social care.*
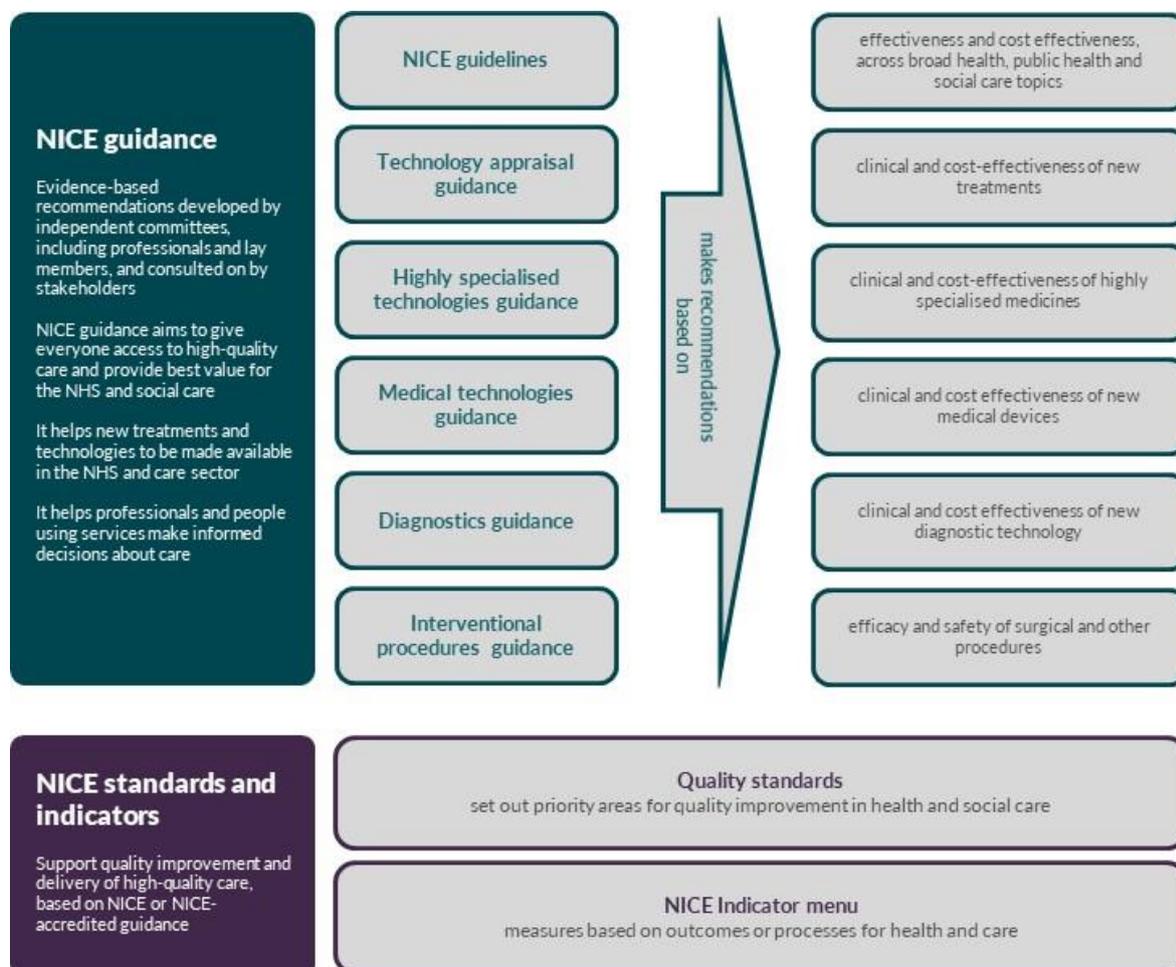
What kind of evidence does NICE currently use to develop guidance?

10. Clinical, social care and public health guidelines have often been based on research findings published in academic journals. These findings may be quantitative or qualitative, and may come from a range of study types, including randomised controlled trials (RCTs), cohort studies, cross-sectional studies and surveys of the views and experiences of people using health and social care services.

11. Synthesis of published evidence may involve meta-analysis of quantitative findings or thematic analysis of qualitative data. All data sources contributing to evidence synthesis are quality appraised using standardised and validated tools.

12. When formal published evidence is not available, NICE has also used evidence from other sources, such as grey literature reports, and expert submissions from health and social care professionals and patients.

13. Guidance developed within our technology-led programmes involves a review of evidence submitted by health technology manufacturers, which covers the evidence they consider relevant to their claims (such as key RCTs), as well as information submitted by consultees, selected clinical experts, NHS commissioning experts and patient experts. The guidance covers the clinical and cost effectiveness of new drugs and highly specialised technologies, the efficacy and safety of interventional procedures, and the evaluation of new medical devices and diagnostic technologies for adoption in the NHS. The Technology Appraisals programme also includes the Cancer Drugs Fund, which supports early managed access to new cancer therapies.

14. We also support new data collection via NHS England's Commissioning through Evaluation programme. This programme supports the creation of new data sources, for example by using established or bespoke registries and linking to other sources of data such as ONS and HES.

15. We currently use a variety of other data sources, including patient registries, audits and electronic health records. Case studies of our existing work using these types of data can be found in the appendix. We recognise the value these data sources have for our work, and plan to amend our processes so that data analysis is considered in more of our programmes. To succeed in this, we will need to work closely with other organisations across the health and social care system and the wider data community.

What broader types of data are available?

16. A wide range of data sources are available. These include formal observational research datasets, as well as collections of data not originally collected for research purposes.

17. The selection of a data source will depend on the review question. Some questions are likely to need quantitative data that reflect the use of interventions in practice. These include questions about treatments and other interventions, or different approaches to delivering public health or social care programmes. On the other hand, questions about preferences of people using services may need to be addressed using text-based data that better capture people's experiences.

18. Some questions may require multiple datasets to be linked together. Linking datasets together offers opportunities to answer questions that would not be possible with a standalone dataset, for example by providing additional fields. Linkages may be made at the point of analysis, or as part of wider programmes

to link different datasets (for example, the creation of Local Health and Care Record Exemplars).

19. Sources that NICE has already used, or could consider using in the future subject to appropriate quality assessment, include:

- primary care databases (for example, Primary Care Mortality Database [PCMD], Clinical Practice Research Datalink [CPRD], the Health Improvement Network [THIN], QResearch)

- secondary care databases (for example, Hospital Episode Statistics [HES])

- other care databases (for example, Community Care Data Set, Mental Health Services Data Set)

- registries containing care and health outcomes data (for example, National Cancer Registration and Analysis Service)

- audits of clinical practice, and registries of the use of medicines, devices and other technologies

- surveillance and monitoring data (for example, data on the uptake of public health preventive interventions, drug safety monitoring data)

- datasets held by local authorities about public health and social care

- patient-reported outcome measure (PROM) data

- quality of life data

- data collected through digital health technologies, including apps, wearables, implantable devices, and the 'internet of things' (consistent with the NICE Evidence standards framework for digital technologies)

- genomic data (for example, UK Biobank)

- reports and outputs produced by health and social care organisations such as the Care Quality Commission and Healthwatch

- grey literature (surveys, reports and datasets that have not been formally published or that have limited distribution)

- qualitative data that represent the views and experiences of people using services, whether captured formally (for example, via academic research studies or by organisations such as healthtalk.org) or informally (via online discussion forums and social media)

- data collected by patient organisations or charities, which may fall into a number of the above categories

- data submitted by experts (which may be quantitative [for example, data collected in a specific context] or qualitative [for example, expert testimony]).

- 'synthetic' data (for example, data that mimics plausible health and care data without representing real individuals)

- data from non-health and social care sources that relates to determinants of health, such as environmental data or social deprivation data.

20. There are existing initiatives in this area, including the creation of digital hubs by Health Data Research UK, Local Health and Care Record Exemplars, the Accelerated Access Collaborative, and academic-led initiatives.

21. Specific challenges exist for social care datasets, for example in terms of the number of datasets available. However, NICE is interested in making more use of social care data, and will do so when this is possible.

22. If RCT data are available for broader use (for example when owners of the data are willing to share the data after the trial is complete), trial populations may act as cohorts for further analysis.

23. Some data sources may be publicly available (open access) and free to use, whereas other sources may only be available with permission from the data owner. In either case, use of data may be subject to specific terms and conditions. Depending on the nature of the project, we may work with external partners who already have access to these resources or we may seek direct access to data ourselves.

24. When there are no suitable UK-based datasets, non-UK data sources may be considered on a case by case basis. For example, with novel treatments or new devices early safety and effectiveness data may only be available from non-UK settings. These data will still need to be generalisable to the UK health and care system to some degree for them to be useable. NICE already has ongoing working relationships with a number of international organisations via our research programmes. However, access to non-UK data sources may depend on the international data-sharing frameworks and legislation that is in place at the time.

25. NICE acknowledges that new sources of potentially relevant data may become available in the future, and will keep abreast of technological and administrative developments to assess whether these may be useful in the development or evaluation of our guidance.

When and why should broader types of data be considered?

26. There are a number of circumstances in which NICE could improve our methods and processes by making more extensive use of data. By doing this, we aim to increase the impact of our products so that they provide greater benefit to people using health and social care services.

27. For many of its products, NICE has used systematic reviews of published experimental or observational evidence, conducted in line with carefully developed protocols. There are also circumstances in which analysis of data has been used to inform decision-making. This section describes a range of situations in which analysis of data could be used in the context of NICE guidance and advice.

Where an evidence gap has been found

28. For some types of NICE guidance, the committee developing the guidance may make research recommendations if a systematic review has identified gaps in the evidence. These research recommendations offer structured questions that, if addressed, may produce evidence that can be used in updates of the guidance. NICE works with the National Institute for Health Research to promote research recommendations and encourage researchers to carry out work that addresses these questions.

29. Some evidence gaps could be addressed with data analysis. We could undertake or commission work that addresses specific evidence gaps, so that recommendations can be made before the guidance is published. Alternatively, this work could inform future updates of the guidance.

To measure the effectiveness and cost-effectiveness of interventions in real-world settings

30. The observed discrepancy between the effects of a health intervention in routine clinical practice (effectiveness) and the effects demonstrated in RCTs (efficacy) is known as the 'efficacy effectiveness gap'.

31. Contextual factors that interact with the effect of an intervention are known as drivers of effectiveness. These factors can contribute to an efficacy effectiveness gap. There are 3 levels of contextual factors:

- the actual use of the intervention (for example, adherence, co-medication, dose/intensity, duration of use)

- patient and disease (for example, age, gender, behavioural factors, baseline risk, genetics, severity of disease, comorbidities)

- healthcare system (for example, implementation, medical practices, screening policies).

32. RCTs may not always reflect these contextual factors, and high-quality observational data may better indicate the expected effectiveness in routine clinical practice. For most decisions relevant to NICE guidance, the expectation is that the efficacy measured in well-designed trials will be greater than the effectiveness derived from analysis of broader types of data. This is because in routine practice characteristics such as age, comorbidities and adherence are more likely to reduce the effectiveness of the intervention compared with trial settings than enhance it. Qualitative data sources may also provide useful contextual information, for example by providing detail on any challenges in delivering an intervention or on the experiences of patients.

To demonstrate comparative effectiveness

33. NICE's assessments require an understanding of interventions in the context of alternative treatments (comparative effectiveness). RCTs remain the optimal design for assessing the comparative effectiveness of interventions. However, robust evidence of comparative effectiveness (either from direct comparisons or indirect/mixed-treatment comparisons) is often not available. This may be because the available evidence does not reflect practice, for example, because the use of an intervention has been optimised since the relevant RCT was conducted. When good-quality data sources are known to be accessible for interventions that lack sufficient comparative evidence to support decision-making, the developer may consider using data analytic techniques (for example, propensity score matching) to explore the effectiveness of the intervention.

To monitor and evaluate intermediate outcomes of interventions

34. RCTs are often not long enough to demonstrate all the outcomes over the full time horizon of interest, especially for treatments for chronic conditions. In these situations, economic analyses often include extrapolation of outcomes over a longer time horizon. Committees considering review questions without economic analyses should still take into account the long-term effects of interventions when relevant. Analysis of data could be used to provide evidence of long-term effects as well as quantifying rare but serious adverse events that may not have been captured because of the limited time horizon or sample size.

35. Similarly, health economic models are often used to extrapolate the effects of interventions beyond the observed data. Examples of this include tying intermediate outcomes to other outcomes of interest, extrapolating the effectiveness of interventions further forward in time and making assumptions about downstream treatments and events. The developer could consider using analysis of data to directly inform or to validate these modelling assumptions.

To establish the characteristics of the population of interest in practice

36. RCTs are often a poor source of information about the characteristics of the true population to which the decision applies because their populations can be highly selective. Broader sources of data can provide more accurate information on the true population of interest, including demographics, disease severity and comorbidities. This may be particularly important when an economic model is being constructed and the composition of the simulated population is an important determinant of expected costs, benefits and harms. For example, trials often provide the best estimates of relative effectiveness, but these estimates should be combined with the most appropriate baseline population/risk in order to estimate absolute effectiveness, which is often the information of interest in economic evaluation.

37. Large-scale data sources may also be used to stratify populations, identifying groups of people who share particular characteristics. This allows evaluations of how effects of treatments can differ amongst particular patient cohorts, and could lead to more tailored or personalised care.

To improve tracking of guidance implementation, uptake and impact

38. Analysis of data can be useful during the surveillance and update of guidance. Evidence generated from data, such as patient registries, can provide valuable insight into use of healthcare resources and implementation of interventions. This may be particularly valuable for recommendations that require a change in service provision, to feed into a 'Learning Health and Social Care System'.

To update guidance more efficiently than we do currently

39. We are currently investigating the automation of some of our horizon-scanning processes, to identify updated sources of evidence more quickly and cost-effectively.

40. Using a broader range of sources of data may enable us to identify when guidance needs a review sooner than scheduled. For example, guidance updates may be triggered by 'early warning signals' relating to the safety of devices and implants and using a wider range of data sources may help us detect these signals. Data reported by the public via reporting systems (or more informally via online forums and social media) may provide signals that are not identified through other channels, though the reliability of these data sources will require further evaluation.

41. The use of broader data sources will not provide any operational efficiency compared with systematic reviews of published evidence. However, new analyses or data collection may improve efficiency when compared with waiting for this information to appear in the published literature.

To better understand the experiences of people using services

42. Published studies do not necessarily capture the experiences of individuals when assessing the efficacy of an intervention or the organisation of a care pathway. Understanding the needs and experiences of people using services is particularly important for social care guidelines, which also often lack other sources of evidence to inform recommendations.

Analytical work in the context of a 'Learning Health and Social Care System'

43. The model of the Learning Health System is based on a continuous loop between practice and evaluation. This loop generates feedback, which can help improve standards of care (figure 2). NICE is interested in how this approach could be used in both the health and social care systems.
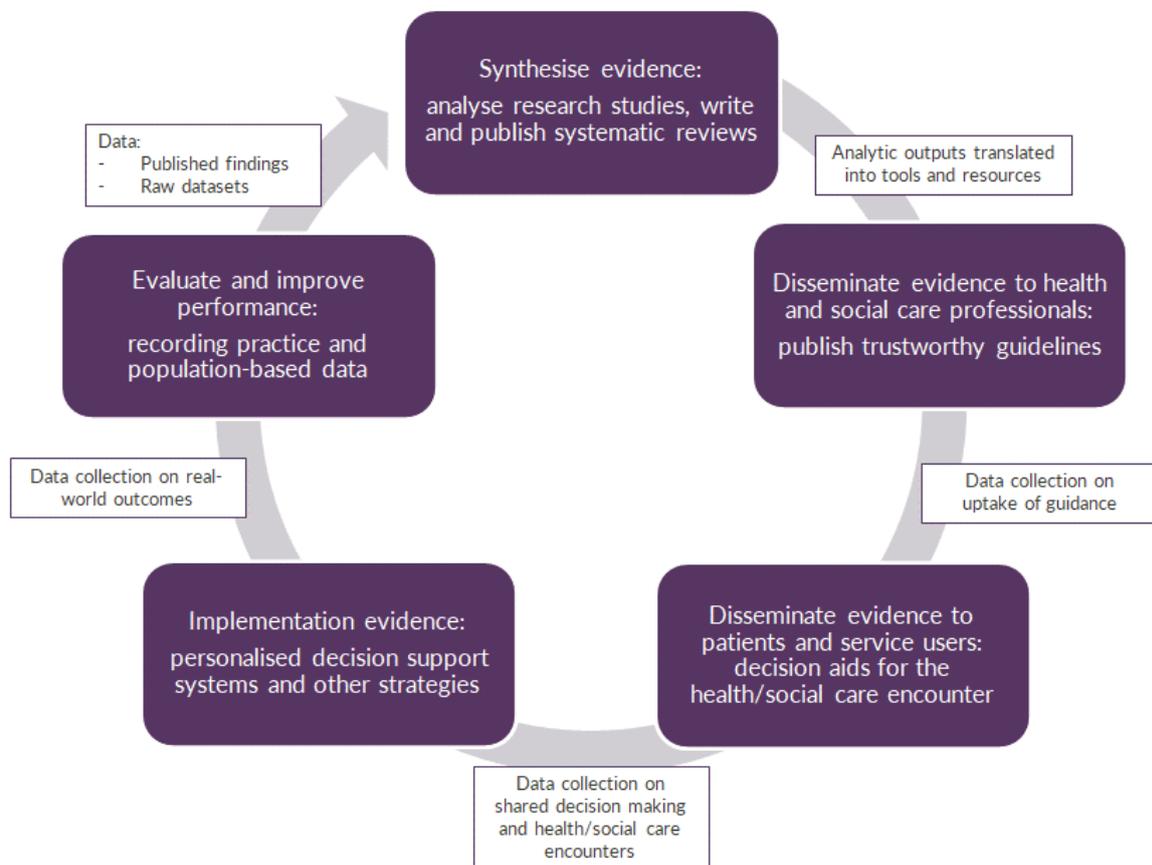
Figure 2: How insights from data can support a Learning Health and Social Care System. *The cycle starts with the synthesis of evidence in order to produce guidance. This guidance is disseminated to health and social care professionals. They share evidence with patients and service users, as well as using it to help make decisions during the health and social care encounter.*

By monitoring implementation there are opportunities to evaluate and improve performance. This evaluation and performance improvement generates further evidence which can feed back into the development of updated guidance, completing the cycle.

44. It is currently challenging to ensure efficient and reliable flows of data and information throughout the system. This requires both a supporting infrastructure and people who are invested in the process at all stages of the cycle.

45. However, there is great potential for NICE's work using broader sources of data to contribute to the learning system, in particular using audit and registry data and long-term intervention follow-up from cohort studies. There are also opportunities in the commissioning and collection of new data, and in the use of novel ways to monitor the uptake and implementation of NICE guidance. Social media data may also be of value, for example in producing early warning signals for harmful drugs or devices.

46. Just as NICE's systematic review work identifies evidence gaps in the scientific literature, it is likely that work involving broader sources of data will highlight areas where data collection is lacking, or where the existing data are of poor quality.

Practical considerations associated with data analytics

Operational considerations

47. There are a number of different operational models NICE could use for projects involving data, whether the data are being used to enhance existing products and services or to develop new ones. In many cases, work involving data will require involvement from third parties. This may be to access particular expertise, or because it is more efficient. For example, NICE may hire external centres to carry out analytical work on an ongoing or ad hoc basis. When work is conducted within NICE, external topic, methods or data experts would be consulted as appropriate. Figure 3 outlines 4 different possible delivery models.

48. The nature of specific projects may determine the choice of delivery model. Factors to consider include:

- whether data are publicly available or require access to be granted
- whether or not NICE has sufficient technical and topic expertise to carry out the project internally
- project delivery timelines
- cost.

Figure 3: Potential delivery models for NICE projects involving data

- NICE receives an analysis carried out by a third party and performs a quality assurance assessment.

- NICE identifies the need for a piece of analysis using data and commissions a third-party organisation to carry out the work. NICE checks the work for quality.

- NICE works in partnership with a third party to identify the need for analysis using data and carry out the work.

- NICE identifies the need for a piece of analysis using data and carries out the work in-house.
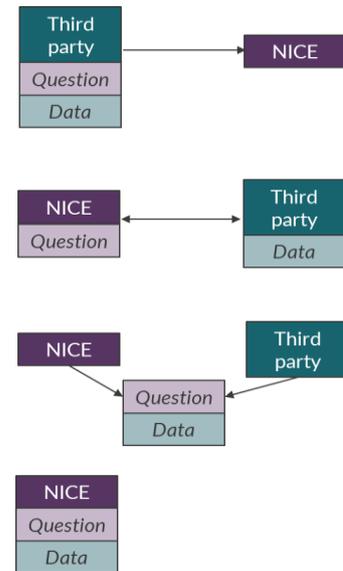
Feasibility and impact

49. Developers (whether they are teams within NICE or other organisations) should consider the practicalities of accessing data, including access costs and the possibility of changes to data ownership. This will need to be done in a timescale that meets the needs of the product under development. Developers will also need to consider whether the data needs processing so that it can be used to answer a review question. This can be a time-intensive part of analysis and require collaboration with subject experts who have knowledge and understanding of the data.

50. Regardless of how the project is organised or where it is done, the same key steps will apply for most larger projects involving anything more complex than descriptive statistics and/or using open access data sources. These steps include:

- defining the project or research question

- obtaining access to the data, fulfilling information governance requirements

- conducting analysis

- quality assurance of data handling and analysis

- reporting findings.

While we are currently streamlining our processes, not all of these steps are under our control or that of our partners, so developers must investigate the likely timelines for a project as early as possible.

51. The potential impact of the analysis on decision-making, including the likely robustness of the results, should also be considered alongside the cost and time needed. Proposals to undertake data analysis will be considered by NICE staff with responsibility for quality assurance and the NICE data and analytics team.

Governance requirements

52. Collection of new data involving people will usually require ethical approval.

53. When data already exist, formal approval may still be needed to use the data for a new purpose (such as research). In such cases a less complex process is followed, but approval is still needed from a research ethics committee.

54. Project development timelines need to factor in this process as a key step in delivery.

55. In addition to the above, NICE will need to ensure that all our data handling and processing complies with the requirements of GDPR legislation.

56. NICE will seek to engage with the public on any information governance or ethical issues arising from the use of data, as part of the broader public discussion on the use of such methods and data sources.

Methodological considerations

57. The recommended processes and methods for incorporating data analytics into NICE products and programmes will be embedded in future methods manuals. Issues include:

    • defining questions that address questions relevant to NICE's remit

    • identifying appropriate sources of data to address these questions

    • evaluating data quality

    • selecting appropriate analytical methods

    • acknowledging the issues of bias and confounding present in observational data sources.

58. We will actively engage with external partners to ensure that the methods used reflect best practice in data analytics and make use of new innovations as appropriate.

59. Quality assurance approaches will be applied both to data sources and the analysis carried out using the data.

60. Data quality should be addressed at both the resource and at the question level:

- The general quality, validity and trustworthiness of a data source should be assessed.

- Additional quality assessment of data is needed for the specific question being considered, and the data items needed to address that question.

For example, a generally well-validated data source may not be suitable for addressing a specific question because of high levels of missing data in key variables, inconsistency of coding, or other issues.

61. NICE has worked with EUnetHTA to develop a tool to assess the quality of registry data. This and continued collaboration with strategic data controllers across UK health bodies will be essential to the control and continued improvement of data quality.

62. Quality assurance processes will be included in future NICE methods manuals and aligned to the principles stemming from the MacPherson review and the 'Aqua Book: guidance on producing quality analysis. The following requirements will be included:

- ensuring that a well-designed analytical protocol has been agreed

- checking whether the analysis reflects what was agreed in the protocol

- assessing whether data handling and cleaning decisions were acceptable, and whether populations, interventions, comparators and outcomes have been correctly defined from the available data

- verifying the code used to perform data cleaning and analysis

- statistical validation of findings as appropriate

- communication of any uncertainty around the outputs of analysis.

Supporting committees and stakeholders to make use of data analytics in developing NICE products

63. We intend to use broader types of data and evidence to supplement existing types of evidence, and to help guidance committees make recommendations. As with other evidence types, data analytic outputs are intended to aid committee decision-making rather than to prescribe a particular decision.

64. While some committees and stakeholders already have experience of using outputs from unpublished research or newly commissioned data, others may need additional support to consider broader types of evidence.

65. NICE has a well-established Public Involvement Programme that already contributes to work in this domain (see case study 4 in appendix A). We will

continue to involve the people using health and care systems in our future work in this area. We will seek their perspectives on ethical issues around using data from these systems, and on other relevant areas (such as identifying which questions are important to the public, patients and carers).

Transparency and reproducibility

66. NICE guidance and advice are produced with a high level of transparency, and new work using broader data sources must maintain these standards. In line with our other guidance development activities, we plan to publish documentation that clearly outlines the questions we are seeking to address, the methods used to generate evidence, and how we will evaluate the quality of that evidence.

67. As part of a commitment to transparency, NICE will aim to share relevant materials that would aid reproducibility when this is not prohibited. There are various reasons why we may be prohibited from sharing data. For example:

- access to certain types of data may restrict how NICE can share specific aspects of a piece of analytical work (for example, it may not be possible to make the raw datasets available if these were commercially sensitive).

- ethical and privacy considerations may prohibit the sharing of some data or outputs because of the need to avoid the risk of identifying individuals. This risk increases as the number of data sources linked together increases.

68. Subject to these challenges, NICE would seek to publish as many of the following items as possible to maximise transparency and reproducibility:

- review question protocols

- methods

- quality assessment of data sources

- details of data cleaning processes

- analysis plans

- any code used as part of the above.

Risks arising from the proposals outlined in this document

69. We have considered a number of potential risks arising from our proposals, across methodological, operational and reputational areas. A summary of these is outlined in Appendix B, and further detail will be provided in future documentation around our methods and delivery proposals.

Summary

70. NICE welcomes the opportunity to continue our well-established practice of translating evidence into practical guidance and advice, and to expand our methods and processes to enable more extensive and effective use of broader sources of data.

71. We acknowledge that there are challenges in expanding our use of data and analytics, but we believe that the potential benefits to health and social care providers and users of their services outweigh the risks. We look forward to engaging with other organisations that have expertise in data analytics, to explore areas of shared interest and work together to improve health and social care. We will seek to reduce barriers and set up frameworks to enable this work to succeed.

References

The Learning Healthcare System: Workshop Summary. Institute of Medicine (US) Roundtable on Evidence-Based Medicine; Olsen LA, Aisner D, McGinnis JM, editors. Washington (DC): National Academies Press (US); 2007.

Appendix A: Case studies of work at NICE

NICE currently uses broader sources of data in several ways. The following are examples of what we have already done.

1. Supporting health economic modelling

We often rely on economic modelling to aid decision-making. To build economic models, good-quality evidence from large patient record databases is commonly used to bridge the gap between efficacy data from trials and the potential application of the technology to NHS practice. For example, in the type 2 diabetes guideline the THIN database (which collates demographic and clinical data in primary care from over 200,000 people with diabetes) was used to establish the baseline characteristics of the population being simulated in the economic model.

In the Parkinson's disease guideline, we used data from the Parkinsonism incidence in north-east Scotland study (PINE). This is a rich dataset describing an incident community cohort with up to 10 years' follow-up. The data was used to estimate the long-term prognosis of people with Parkinson's disease and the ways that different treatments could be expected to influence the natural history of the disease.

2. Supporting new data collection to support NHS England Specialised Commissioning policy decisions where trial data are not yet available or are not generalisable

NHS England's Commissioning through Evaluation programme gives selected groups of patients access to promising treatments that are not yet funded by the NHS. We work with external partner organisations who collect data on clinical outcomes and patient experiences. Analyses based on these data then contribute to a formal evaluation programme.

To date, 3 treatments provided through Commissioning through Evaluation schemes are now routinely commissioned in the NHS in England:

- left atrial appendage occlusion to prevent stroke in patients with atrial fibrillation

- selective dorsal rhizotomy to reduce spasticity in children with cerebral palsy

- selective internal radiation therapy for unresectable colorectal cancer metastases

A further 2 treatments have decisions pending (patent foramen ovale closure to prevent recurrent stroke, and MitraClip for mitral regurgitation).

3. Providing evidence for health technology appraisal reviews

The committees that advise NICE on technology appraisals and highly specialised technologies are able to recommend that technologies are used subject to a managed access agreement (MAA). These typically include a specified period of data collection to inform a review of the guidance. Data collection usually includes elements of observational data. There are currently 30 MAAs collecting observational data across the Technology Appraisals and Highly Specialised Technologies Programmes. For technologies that are recommended as part of the Cancer Drugs Fund, we work in partnership with NHS England and Public Health England to collect observational data via the Systemic Anti-Cancer Therapy Dataset to inform reviews of the guidance.

In 2018, NICE recommended one of the first [CAR-T cell therapies for use in the CDF](#). As part of the [managed access agreement](#), observational data are being collected using the Systemic Anti-Cancer Therapy Dataset, Hospital Episode Statistics and the UK bone marrow transplant registry. We expect that these data will help resolve uncertainty identified by the NICE committee developing the guidance, when it is reviewed in 2022.

4. Involving the public in our work by collecting data on the experiences of people using health and social care services

Our Public Involvement Programme (PIP) has a number of ways to support the use of data covering the experiences of people using health and social care services.

Our Interventional Procedures and Medical Technologies programmes sometimes collect new data on patient experiences through surveys of people using health and social care services. People are invited to participate by the professionals involved in their care, or by collaborating patient organisations. The data collected in these surveys are processed and collated by the PIP and passed on to the committees that develop NICE guidance.

Patient organisations are invited to submit information about the experiences of their members, to support our Technology Appraisals, Highly Specialised Technologies and Interventional Procedures Programmes. In some cases, submissions from patient experts will also be sought.

Our health and social care guidelines use grey literature for some guideline topics, to inform scoping or guideline development. This may include surveys conducted by organisations representing users of health and social care services (or their carers), data from online forums, and websites such as healthtalk.org.

We will involve members of the public and users of health and social care services in any work based on the proposals in this document.

5. Monitoring how NICE guidance is being used in practice

We want to understand how the recommendations in our guidance are used in health and social care. There are various sources of data that can help with this. Our Adoption and Impact team routinely searches data sources such as THIN (primary care), Hospital Episode Statistics (secondary care) and prescribing data sources such as ePACT2 (primary care) and HPAI (secondary care).

In a recent example, the Adoption and Impact team were asked to look at Hospital Episode Statistics activity data on people admitted to hospital who have a risk assessment for venous thromboembolism or bleeding. Diagnosis, age, sex and length of hospital stay were considered, and the team found that length of hospital stay made a difference to the cost to the NHS of using different risk assessment tools. In the next guideline update, we made a recommendation to address this.

6. Ensuring that guidelines are updated when new evidence emerges

Our Surveillance team checks that our published guidelines are up to date. This is done by searching for new evidence that could contradict, reinforce or clarify guideline recommendations. In addition to evidence, 'intelligence gathering' plays an important part in surveillance decision-making. Data sources that the surveillance team use include:

- antimicrobial prescribing data collected by Public Health England, to evaluate the impact on our antimicrobial prescribing guidelines

- data indicating how guidelines are being used in practice, for example uptake data for our osteoporosis guideline collected by a Royal College of Physicians audit

- MHRA drug safety monitoring data, which can be incorporated into surveillance reviews of guidelines with recommendations on medicines

- prescribing data for medicines recommended in our guidelines.

Appendix B: Risk register

NICE has identified the following potential risks arising from its proposals, and we will seek to mitigate these as we further develop our data and analytics work

Category	Risk	Mitigation
Information governance	Risks of data breaches or misuse of sensitive data.	Ensure that NICE and partner organisations are fully compliant with regulations and best practice in data handling, reporting and security. Ensure that patient and public confidentiality is maintained.
Quality and validity	Risk that novel data analytic work conducted or commissioned by NICE may generate outputs where quality and validity is a concern.	NICE to engage with partners to develop shared understanding of methods and processes before and during projects. NICE to clearly communicate its needs and standards, including: - clear definition of questions to be addressed with data analytics - advice on data cleaning and management - ensuring that the most appropriate methods are used to address the question - communicating uncertainties and limitations in any data or analysis - demonstrating a strong commitment to transparency and reproducibility (for example by making data cleaning and analysis code available for review) - supporting committees in interpreting results, to support the process of developing guidance and recommendations.
Timely delivery	Risk of failing to deliver projects within the normal parameters of NICE's product development timelines.	NICE to conduct feasibility assessments at the outset and during a project to ensure that the proposed project has feasible delivery. Feasibility assessments to include both data access arrangements and availability of supporting resources. NICE to work to reduce barriers to timely delivery, including:

		- developing a clear project brief at outset, in conjunction with relevant stakeholders - engaging in ongoing learning about factors that may affect the feasibility and timeliness of projects (such as the data quality of specific resources) - streamlining processes as far as possible with respect to accessing data and building relationships with relevant stakeholders and experts.
Reputational risks	Although NICE already carries out or commissions work in this area, this is not widely known. Health and social care practitioners may have questions about the direct use of data in NICE's products and programmes.	NICE will raise visibility of its established work in this area. Where new approaches to the use of data are introduced into specific products, the strengths and limitations will be made clear. NICE will communicate to health and social care practitioners that use of broader sources of data will not replace NICE's existing methods, but rather will supplement these where value can be added through the extended use of data and analytics. Data analytics outputs will, as with other types of evidence NICE considers, be available to support decision-making, but not prescribe a specific recommendation.
Public trust	Risk that public perception of NICE's proposed expansion of its use of data and analytics undermines public trust in the organisation or its outputs.	NICE to engage in the wider conversation around developing and maintaining public trust regarding the reuse of data from health and social care settings. NICE to ensure that patient and public data are used in an appropriate manner to deliver genuine benefit back to people using health and social care services.
Quality and validity	Risk of the corollary effect of NICE accepting lower quality data leading to lower standards in general.	NICE will communicate to health and social care practitioners that use of broader sources of data will not replace NICE's existing methods but rather will supplement these where value can be added through the extended use of data and analytics. Data analytics outputs will, as with other types of evidence NICE considers, be available to support decision-making, but not prescribe a specific recommendation.

Miscellaneous	Risk of new technology and methods arising that lead to unforeseen issues.	NICE will engage with others at the forefront of development and keep abreast of new technologies/data/methods in order that they can be utilised and/or accounted for.
Miscellaneous	Risk of the possible additional burden of data collection.	NICE will aim to ensure that any additional data collection requirements are minimised and proportionate.
Quality and validity	Risk that bias may exist in data and subsequently impact analysis.	NICE will create and use a methodological approach to data and analysis that identifies and minimises bias. Any deficiencies should be explicitly addressed.
Quality and validity	Risk of conflicts of interest affecting data acquisition and use.	NICE will work to ensure that the provenance of data it uses of is known and acknowledge and attempt to adjust for any potential bias that may arise.
Quality and validity	Risk of disagreement between different data sources.	NICE will seek to use best practice or where appropriate contribute to the development of new methods as needed, to resolve differences between evidence sources in the context of decision making.
Miscellaneous	Risk that complex terminology in this field will impair understanding of this work by stakeholders (including the public).	NICE will define key terminology via its glossary and aim to communicate outputs in plain English.
Miscellaneous	Risk that semantic variation could lead to poor communication between stakeholders and collaborators from different technical fields.	NICE will make use of an ontology that attempts to encompass the data science field in the context of evidence-based guidance development. This will minimise semantic variation.