BIA submission:
The scope of qualifying expenditures for R&D Tax Credits

October 2020

UK BioIndustry Association

Introduction

The small and large business R&D Tax Credit schemes are often cited by BIA members as the most valuable form of innovation support the Government provides. Tax credits provide a minimal-bureaucracy system that incentivizes and amplifies companies' own investment in R&D. Ensuring the scheme remains internationally competitive, up-to-date with developments in the economy and how research is conducted in the 21st Century, and functioning as intended, is critical to maintaining the UK's attractive fiscal environment for R&D investment and achieving the Government's stated ambition to raise R&D investment to 2.4% of GDP and make the UK the global hub for life sciences.

Governments around the world are keen to attract high-value, R&D-intensive industries to their shores. These businesses generate many spill-over effects, including new technologies, productivity gains from process innovation, down-stream manufacturing jobs, and the co-locating of regulatory, legal, and financial services. Moreover, the COVID-19 pandemic has brought into sharp focus the value of a vibrant domestic life sciences sector underpinned by a competitive business and fiscal environment.

The Government's focus on raising public and private R&D investment and strengthening the life sciences sector are critical to the UK's economic and social recovery from COVID-19. Tax expenditures such as R&D tax credits are an important part of the policy mix required to achieve these objectives. Maintaining the international competitiveness of the UK's tax regime will ensure the UK remains a global destination for life sciences investment and so we welcome Government commitments to consult on the inclusion of expenditure on data and cloud computing as an eligible cost for R&D tax credits, ensuring the regime is fit for science in the 21st Century. This change will bring the UK regime into line with our international competitors, notably America.

The central importance of data and cloud computing to life sciences R&D

Today we are witnessing the simultaneous maturing of numerous breakthrough technologies – genomics, nanotechnology, sensors and the Internet of Things, big data and advanced analytics, artificial intelligence (AI) and robotics – that is unprecedented in human history. Within life sciences R&D, the convergence of these presents the opportunity to ensure better outcomes for patients via targeted therapies, significantly reduce the cost of drug development, and accelerate cycle times to get treatments to patients faster. In healthcare more broadly, the opportunity is better care, delivered at the right time, for lower cost and with fewer mistakes and side effects.

These opportunities can help address one of the fundamental challenges to humankind, which faces every economy across the globe: how to sustain healthcare systems serving growing and aging populations that demand safer and more effective treatments.

The business opportunity is also huge. As we have seen in social media and advertising, the companies that best harness the power of data are rapidly scalable, becoming significant engines of economic growth and

job creation. Economies that support their businesses to access and use data will benefit from the growth of these companies and their citizens will benefit from the products they develop.

A prime example of this is the recent clinical trial result[1] that showed the combination of two drugs - baricitinib plus remdesivir – was more effective at treating serious COVID-19 disease than remdesivir alone. The combination was identified using UK-based BenevolentAI's knowledge graph, which is a large repository of structured medical information, including numerous connections extracted from scientific literature by machine learning. This was used to search for approved drugs that could help, focusing on those that might block the viral infection process. This identified baricitinib on the 15 February, mere weeks after the structure of the COVID-19 vaccine was mapped. This demonstrates the power of data for drug discovery.

Answers to consultation questions

Question 1a - Are there uses of data that contribute to R&D but which do not currently attract relief through the RDEC and SME schemes? Please provide examples to support your response

The BIA was grateful to have had the opportunity to arrange two roundtables with HMT and HMRC in which our members presented the various ways in which they use data in their R&D projects, which, on the whole, are not deemed eligible for R&D tax credits by HMRC.

These uses included activities such as:

- Analysing and modelling data generated from laboratory experiments to understand biological processes and drug interactions as part of the drug discovery and development process
- Using natural language processing (NLP) of unstructured data sources (patient records, academic papers, patents etc.) to extract information and new insights
- Applying machine learning (ML) and artificial intelligence (AI) to drug molecular structure datasets to identify potential drugs and/or improve drug function
- Mining genomic datasets combined with patient records to identify links between genes and the manifestation of disease, allowing new treatments to be developed and non-therapeutic interventions to mitigate illness
- Analysing health records to identify commonalities in patients' responses to treatments in order to better target treatments to patient populations that will respond positively, thus reducing wasteful spending and unnecessary exposure to side effects
- Analysing prescription data and patient records to improve efficiency in healthcare delivery systems
- Analysis of supply chain data and modelling to improve supply chain resilience, bring down costs and deliver safer medicines to patients

These activities generally require the purchase of data from third parties; the experience of our members is that they are not able to claim for this expenditure under the R&D Tax Credits regime at present. These costs are often high, especially for large, detailed and homogenised datasets, which are the most valuable

[1] https://www.benevolent.com/news/niaid-releases-additional-data-validating-benevolentais-hypothesis-for-baricitinib-as-a-treatment-for-covid-19-patients-in-large-scale-randomised-control-trial

for research purposes. Moreover, for SMEs, accessing high-quality data is often cited by BIA members as a barrier to expanding their R&D programmes and maximising the gains from them.

As discussed further in answer to the next question, raw datasets are only the feedstock and considerable cleaning, processing, collating and manipulating of data is required to transform them into usable datasets for use in R&D. It is also through this process that true value in the data is created. AI, ML, and NLP is used to consume, summarise, interrogate, link and derive insights from structured and unstructured data at a scale previously unimaginable. The key to unlocking insight is combining all information sources together and deriving insight and new ideas from what is known. These transforming processes, which often involve cloud-based storage and computing power, are costly and, again, our members on the whole report that they cannot claim for them under the R&D Tax Credit scheme.

Question 1b - To what extent are datasets employed in the R&D process consumed? To what extent do they retain value? Please provide examples to support your response.

It is often said that data is the new oil. Oil is valuable only for what it can be transformed into – hydrocarbons are broken down, purified and combined with other molecules to create myriad valuable products that can be used in many different ways. The same is true for data.

The intellectual property of companies involved in the use of data in R&D, and therefore the basis of their value, is in the creativity and technology that they have developed to transform data into something valuable, which can be used for R&D. Through this transformation, any fixed dataset in effect provides a one-time benefit through the insights it affords. Accordingly, most current applications rely of datasets that are constantly updated and expanded. In general, raw datasets only retain value for a certain period of time and need to be replenished with new data or updated in other ways to remain relevant.

Data is an essential feedstock for R&D and there is an obvious gap in the legislation and guidance such that this cannot currently be included. This is why confirmation through a policy change by HMT is necessary to ensure the scheme functions correctly and fairly for all companies, and that the scheme rightly supports UK-based companies to invest in data-based R&D and foreign companies to base that valuable R&D activity using data in the UK. We are concerned that efforts to accommodate data into the existing definition of consumables may result in uncertainty and certain valid costs potentially being excluded. Requiring companies to demonstrate that the data has been consumed would create uncertainty and an unnecessary burden on businesses and HMRC inspectors. Data should instead be thought of as a separate category of cost.

Question 2a - Do you already claim for software costs under the current definition? If so, what was your experience of separating out the R&D specific costs for the purposes of the claim?

Software remains an important cost in R&D and the rules and their application under CIRD82500 are understood by claimants without any difficulty. However, software is also increasingly being leased and accessed via the cloud, especially where large data R&D is concerned. Both data and access to cloud computing have been explicitly excluded by HMRC and given the evolution of R&D methodologies as set out above, from a policy perspective, we cannot see why a distinction should be made.

Question 2b - Are there any software costs that currently qualify for R&D tax credits, that could be limited or excluded from relief without materially affecting R&D projects? Please provide examples to support your response.

To the extent that software costs are qualifying under the current BEIS guidelines, and the software is employed in R&D, we don't think it is reasonable to exclude anything that is currently allowable.

Question 2c - Are there any software costs, partially or wholly for R&D purposes, that do not currently qualify for R&D tax credits, that should be if the regime is to better reflect the nature of modern R&D? Please provide examples to support your response on whether these costs could be separated out straightforwardly.

Cloud computing licenses cover the software (or the platform on which companies code their own software), data storage, computing power and associated support services. As described above, data is heavily manipulated throughout the R&D process requiring great amounts of storage and computing power. These license costs can therefore by very expensive relatively speaking for SMEs, but not currently relievable. Expenditure on these activities should be explicitly recognised as eligible for R&D Tax Credits.

The providers of these services understand the need to separate out project costs, either by R&D vs commercial uses or by other definitions, so it would be easy to accurately and reliably report R&D-attributed costs in R&D Tax Credit claims.

Question 3a - What experience do you have of claiming R&D tax credits in other jurisdictions, where expenditures pertain to data or cloud computing?

BIA members report that many other jurisdictions, notably the US, permit data and cloud computing costs within their R&D incentive regimes as standard. The US is notable because it is the primary competitor for the UK in terms of attracting life science R&D businesses. The UK is the clear leading European destination for life sciences – over a third of all new biotechs created in Europe since the start of 2012 have been in the UK[2] - and globally second only to the US clusters of Massachusetts and California for total venture capital raised for biotech[3]. We must therefore ensure our regime is competitive with the US, which, due to Silicon Valley and its vibrant venture capital community, already has a strong draw on internationally-mobile data-enabled life science start-ups and scale-ups. Maintaining a fit-for-purpose tax regime will also ensure the UK is the first destination for European entrepreneurs when looking where to base their businesses. It could also in part counter the negative effects of losing access to European innovation funding and recent detrimental changes to Entrepreneurs' Relief.

Question 3b - What evidence can you provide that a scope expansion in these areas would drive you to make additional investments in research and development?

Increasing the amount businesses can claim through R&D Tax Credits has been shown to increase R&D expenditure. This was shown in a quasi-experimental study by the Saïd Business School at the University of Oxford[4]. Before 2008, SMEs were identified as having 250 employees or less, but in 2008 this limit was raised to 500. This provided a cohort of companies (with a headcount of between 250 and 500) that overnight

[2] McKinsey & Company (2019), Biotech in Europe: Scaling Innovation: https://live-biocentury.pantheonsite.io/sites/default/files/Biotech-in-Europe-Scaling-Innovation-McKinsey-BioCentury-Report-BioEquity-2019.pdf

[3] BIA (2020), *Global and growing: UK biotech financing in 2019:* https://www.bioindustry.org/resource-listing/global-and-growing---uk-biotech-financing-in-2019.html

[4] Irem Guceri and Li Liu, Effectiveness of Fiscal Incentives for R&D:Quasi-experimental Evidence, American Economic Journal: Economic Policy 2019, 11(1): 266–291 https://pubs.aeaweb.org/doi/pdf/10.1257/pol.20170403

became eligible for the more generous SME scheme. In theory, nothing else should have changed, meaning the change in scheme was an isolated variable.

Comparing the newly-classified SMEs on the more generous scheme to the companies that continued to be classed as large, the researchers found that those now in receipt of a more generous R&D tax credit increased their R&D investment by 33%. The more generous R&D tax credit reduced the cost of doing R&D for the company by 22%, so the increase of 33% investment represents incentivised behaviour change, not just reallocation of funds. Furthermore, companies invested £1 for every £1 foregone in tax to the Exchequer, so the policy was cost neutral. R&D-intensive companies and young firms responded most strongly to the tax change. We therefore believe there is strong evidence that extending eligible costs to include data and cloud computing would lead to increased R&D activity with the associated productivity gains, job creation, and spill-overs, as well as leading to more tax-paying innovative businesses in the UK.

Question 4 - Would changes to the R&D tax relief rules in the areas outlined above lead to any change in the commercial relationships between companies, insofar as expenditure is outsourced to a third-party provider?

We do not believe it would, but it is conceivable that, by lowering the costs of the in-house data transformation processes described above, the changes in this consultation could incentivise more in-house activity and less outsourcing of those activities to third parties.

For any further information on the contents of this submission please contact Dr Martin Turner, Head of Policy and Public Affairs, by emailing mturner@bioindustry.org