igniting our potential

Health Research Analysis Forum (HRAF)

The UKCRC Health Research Analysis Forum (HRAF)

The Health Research Classification System (HRCS) is a standard classification approach for the analysis of health research portfolios, developed by the UKCRC partners in 2004. To date the UKCRC has conducted four major analyses of health research using this approach (The *UK Health Research Analyses* in 2004/05, 2009/10 and 2014, and *From Donation To Innovation* (2007)[1]), which have provided a regular snapshot of health research funding across the UK. Such information has had a major impact by providing a unique, unified assessment of UK health R&D to identify gaps and opportunities and a baseline for strategic decision making.

The HRAF (for membership, see annex 1) was established at the end of 2009 to:

- Ensure the continued effective use of the HRCS by UK stakeholders
- Oversee the conduct of any future joint health research analyses
- Respond to future national or international initiatives in this area
- Take responsibility for the HRCS Online website and the databases created for the original UKCRC joint health research analyses

Progress update

1. Chair of the HRAF

The HRAF propose a transfer of chair duties from Dr Ian Viney to Dr James Carter, with effect from June 2018. Jim is an Evaluation Officer in the MRC, was project manager for the last analysis and has continued as primary point of contact for the HRCS website and HRAF business since 2015. The MRC expects to continue to convene the HRAF meetings, provide the secretariat, administer the HRCS website and remains committed to delivering the *UK Health Research Analyses*.

2. Analysis of UK health research portfolios

2.1. Publication of the *UK Health Research Analysis 2014*.

The most recent report in the *UK Health Research Analysis* series was published in August 2015[2], based on 2014 portfolio information from 64 public and charitable funders of health research. This report represented a significant increase in participation from the 12 HRAF funders featured in the original reports, and included 17,000 awards and over £3bn in funding assessed.

The report also assessed changes in funding across the ten years since the first report. Analysis showed that while overall investment in UK health research has increased, the last five years

---

[1] All four reports are available to download as PDFs: https://hrcsonline.net/reports/analysis-reports/
[2] UKCRC (2015) *UK Health Research Analysis 2014.* ISBN: 978-0-903730-20-4 (web-friendly PDF)

showed little overall growth in real terms[3]. In addition, this analysis showed that while fundamental discovery research still accounts for half of all health R&D funding, the proportion of funding in translational research had increased[4].

2.2. Automated coding

The most resource-intensive aspect of analysing health research portfolios using the HRCS is manually coding awards. Since 2012 HRAF has examined the opportunities to automate the coding process. Initial work with Elsevier, using data from 2004/05 and 2009/10 reports, was unsuccessful. However, recently UberResearch Ltd. has used the 2014 report data to develop an approach that has had promising results. The auto-coding approach is now available via UberResearch Ltd.'s *Dimensions* system, which includes award data from publicly available funding sources (e.g. Gateway to Research, Europe PMC, NIH RePORTER). Having the HRCS built into systems marketed globally is expected to increase the use of the approach for portfolio analysis by funders internationally.

In addition to more rapid data collection and coding, access to *Dimensions* raises the opportunity to:
- Identify health-relevant awards from participating funders not exclusively funding biomedical research (e.g. the rest of UKRI not part of HRAF)
- Inclusion of publicly available data from UK funders that have not participated in previous UK health research analyses (e.g. Austica, Bloodwise, DEFRA)[5]
- Access to data from international funders supporting UK health research (e.g. EC/EU)

The drawbacks of this approach include:
- The costs of access to *Dimensions* is too high for all funders to subscribe to
- While gains can be made by improving the quality and availability of data published to existing public databases (from which *Dimensions* extracts data), not all data will be accessible via this system
- Some funders consider manual coding as optimal, and are unwilling to change practices at this time

At present, only half of the 64 funders that participated in the 2014 analysis have portfolio data available on *Dimensions*, which is broadly reflective of individual funder policies regarding publishing award data. As a result of this - based on awards submitted for the last analysis - we estimate that only 60% of awards potentially suitable for inclusion in future analyses are available on this system, and only 53% are successfully auto-coded[6].

Assessment of the automated coding approach suggests that it has similar accuracy to manual coding on an aggregate level, which combined with the significant efficiency benefits of auto-coding has led some funders to already fully adopt this process for award classification. This means that automated coding will be part of any future UK analysis and therefore steps must be taken to account for introduction of this new methodology when comparing between funders and across previous reports.

---

[3] Assessment of compound annual growth rate showed a 8.2% increase in funding between 2004 to 2009, but only 1.4% between 2009 and 2014.

[4] In assessment of HRCS Research Activities, proportion of spending in *Underpinning* and *Aetiology* fell by 9.7% and 5.2%, respectively between 2004 to 2014, with a corresponding increase in research classified as *Prevention* (+2.9%), *Detection and Diagnosis* (+4.7%) and *Treatment Development* (+3.5%).

[5] Presently there are data from seven public/charitable UK funders who have not yet participated in these analyses which could be included.

[6] Note that the *Dimensions* HRCS coding algorithm is based on Boolean text searches with boost terms, with a confidence threshold scoring applied to minimise false positives. Awards not coded either have insufficient published data (e.g. truncated, redacted or missing abstracts) to allow the algorithm to code above this threshold or are genuinely not health relevant.

2.3. Planning for the next UK Health Research Analysis

2.3.1. A continued need to summarise expenditure on UK Health Research

The *UK Health Research Analysis* series has become a vital source of national data on health research funding. While advances have been made in the public availability of research funding data[7], this report series remains the only comprehensive source of UK public and charitable funding data endorsed by the participating funders and showing trends over time. The last analysis in 2014 was successful in reducing reporting costs despite a substantial increase in participation and data.

The HRAF are committed to continue the trend towards improved availability, collection and publication of portfolio data, with specific aims for the next *UK Health Research Analysis* being:
  i.   Further extend the time series we have on health relevant research in the UK
  ii. Extend comprehensiveness through increasing funder participation
  iii. Maintaining data quality through improving coding consistency
  iv. Advance data management and analysis to allow for reduce costs and allow more timely reporting
  v.   Continue to make data collected to be as openly accessible as possible

2.3.2. Cost of analysis

The shared cost of the coding and analysis for the last analysis was estimated at £81,000 (£4.75 per award, versus £7.54 per award for the previous analysis in 2009)[8]. However much of this cost relates to on-going commitment to routine HRCS coding of award data, rather than specifically the *UK Health Research Analysis 2014*. As such our analysis costs are off-set by use of these data in other queries or reports. We would estimate that if coding costs were excluded, the dedicated *UK Health Research Analysis* reporting resourcing cost (i.e. project management, data collation, analysis and writing) would be less then £20,000, with this cost spread across the 12 HRAF members and covered through existing evaluation/reporting budgets.

2.3.3. Project management for data collection, analysis and report writing

In April 2018, the MRC and AMRC continued and extended their agreement to collaborate on the impact of medical research. The agreement includes work to improve the analysis of portfolio data, as well as the collection of output information via Researchfish® and analysis of long-term impact from research. We hope that this will help encourage AMRC members to contribute to the proposed fourth *UK Health Research Analysis*.

Use of UberResearch Ltd.'s *Dimensions* system is will facilitate collection of data from both participating and new funders to the analysis (see section 2.2 for more details). We will also seek to include additional funders through direct discussions, as per the 2014 analysis.

This will still require dedicated project management to ensure all data required for the analysis is compiled from a mix of accessing *Dimensions* and obtaining data directly from funders. As with the last two reports, the MRC would take the lead as main project manager for the proposed analysis, using a similar framework to the 2014 analysis. However, given the use of automated coding by some funders, the next *UK Health Research Analysis* will require some additional analysis to account for the mix between manual and automated coding used.

2.3.4. Timing for the next analysis

---

[7] Examples include NIH's World RePORT, G-FINDER for neglected diseases and the WHO Global Observatory for Health R&D.
[8] See Carter et al. (2016).

The HRAF has conducted an assessment of award duration and changes in portfolio classification over time. The average award length across all funders is 3.5 years, however up to 20% of awards have a duration of five years or more. Assessing the trends within HRCS classified data suggests that four years would be the minimum period that would allow for meaningful changes to occur in the portfolio.

The HRAF suggests that data for *UK Health Research Analysis* remains an important, publicly-available resource both for participating organisations and as an evidence base for wider reporting / decision making. **Given the above criteria, the HRAF recommends that the next UK Health Research Analysis is conducted in 2019 based on 2018 calendar year data.** This report will include a mix of both automated and manually coded award data, but with appropriate use of *Dimensions* data will allow more rapid collection of data from participating organisations and widening participation to additional UK and international funders that fund research within the UK.

3. Review of the HRCS and update to the website

Since the publication of the last analysis in 2015, the HRAF has conducted a qualitative and quantitative review of the HRCS to confirm the system remains fit for purpose, identify areas of coding inconsistency, and prepare updates to the relevant guidance. This process has now been completed and HRAF has agreed that while the overall structure of the HRCS does remain fit for purpose, a total of 42 topics required updates to improve overall coder consistency and/or reflect new areas of research[9].

In addition, the HRAF have undertaken a redevelopment of the current HRCS website. While this work was initially prompted by the previous site leaving active development support (increasing the risk of cyberattack), this has allowed for design changes to be made to better reflect the needs of website users and incorporate post-review guidance updates. The redevelopment costs of £15,600 were shared across HRAF funders and supported by the MRC web development team. The new website[10] launched in February 2018 and feedback so far has been overwhelmingly positive. A formal survey of users will be conducted later in 2018.

Board members are asked to

- Note the change of chair and progress on work from HRAF thus far
- Note the HRAF's proposal for a fourth *UK Health Research Analysis* in 2019.
- To comment / advice on current usage and proposed future work involving the HRCS

---

[9] The proposed guidance updates are summarised here.
[10] The new website is available via the same URL as the previous: https://hrcsonline.net/

Annex 1 – Health Research Analysis Forum (HRAF) membership

MRC, UKRI (chair)	Ian Viney (to June 2018) Jim Carter (from June 2018)
ABPI	Ali Hansford
AMRC	Liz Philpots / Jocelyn Leblanc
Arthritis Research UK	Tammy Cheng
BBSRC, UKRI	Beverley Thomas / Kate Chisholm
British Heart Foundation	Shannon Amoils / Subreena Simrick
Cancer Research UK	Lynne Davies
Chief Scientist Office, Scotland	Julie Simpson
EPSRC, UKRI	Ben Ryan
ESRC, UKRI	Joy Todd
Health and Care Research Wales, Welsh Government	Michael Bowdery / Marc Boggett
Health & Social Care Services, Northern Ireland	Nicky Armstrong
NIHR	Jola Sysak
UKCRC	Sarah Qureshi
Wellcome Trust	Jonathan Best