Igniting our potential

Health Research Classification System: Coder Training Course 2025

Dr. Jim Carter, MRC

v5.2: April 2025

Structure of the training session

- Background (~15m)
  - Where the HRCS came from
  - What impact it has had
  - Who is using it now and why

  "Why should I have to classify award data?"

- Understanding the system (~45m)
  - How the HRCS is structured
  - What it can and cannot do
  - How to use it in practice

  "How do I classify awards?"

- ~~Coding exercise (45mins)~~

Igniting our potential

Health Research Classification System

Background

Origins of the HRCS: UKCRC

Working in partnership. Changing cultures. Igniting our potential

1st UK Health Research Analysis (2004/05)

- Aims
  - Coherent, national approach to funding health research
  - Evidence base of current funding patterns
- Participating Organisations
  - 4 Health Departments
    - England, Scotland, Wales, Northern Ireland
  - 4 Research Councils
    - MRC, ESRC, EPSRC, BBSRC (health relevant data)
  - 4 largest charities
    - Wellcome, CRUK, BHF, Versus Arthritis*
- Designed a bespoke **Health Research Classification System (HRCS)**
  - Research Activity Codes – types of research
  - Health Categories – areas of health and disease

** Part of AMRC's Donation to Innovation additional report, participating fully since 2009/10*

Impact of the first analysis and HRCS

- Reports widely distributed
- Evidence base used by
  - Individual funders
  - Joint planning and coordination
  - Research community
- Classification system adopted by many UK government and charity funders
- Primary data source for Cooksey Review
  - Office for Strategic Coordination of Health Research (OSCHR)

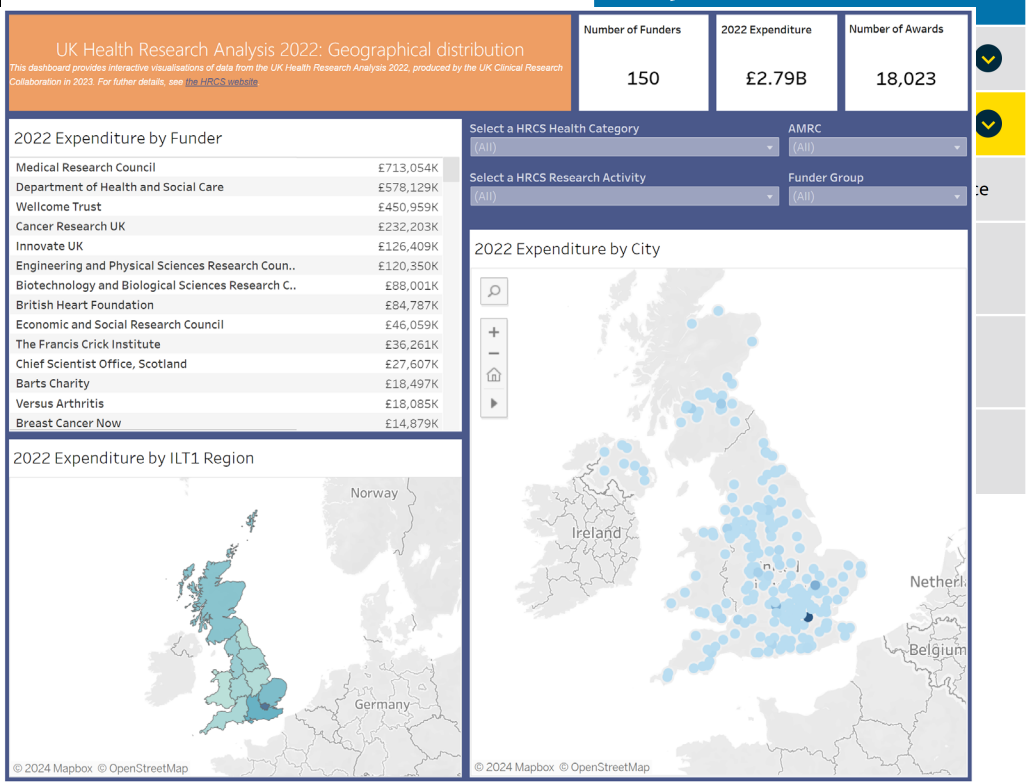
Aims of UKHRA (2009-2022)

- Improve speed of report production
  - Greater availability of data in more easily extracted format
  - The majority of awards already coded
- Increase number of participating organisations
  - More medium/small charities and other public funders
- Improve consistency and quality of coding
  - Review HRCS, update guidance and standardise approach
- Make the data openly accessible
  - Considerable interest in further use of UK HRA data
  - Pre-agreement of HRAF members to release full public dataset
- Expand the report to provide additional analyses
  - Infrastructure Assessment (~£1bn spend pa)
  - Estimation of Total UK Health Research Expenditure

Accessing the reports and data (HRCS website)

Visit the [UKHRA22 Homepage](#) for…

- Report PDF
- Public Dataset
- Tableau dashboard for quick analyses



HRCS ONLINE

Home | Getting started | Health categories | Research activities | Guidance | Reports

UKCRC Health Research Classification System

Search

Home > Reports > Analysis reports > UK Health Research Analysis 2022

UK Health Research Analysis 2022

The fifth in our UK analysis series is officially published

The *UK Health Research Analysis 2022* – the fifth iteration of our nationwide health research landscaping report series – has been published.

Begun in 2023, the *UK Health Research Analysis 2022* provides the most detailed overview yet of UK health research funding from all public sectors, including the governments of the four nations of the UK, charities, societies, and professional bodies.

The analysis consists of more than 23,500 health and biomedical research awards from 173 organisations accounting for an estimated expenditure of just over £5.0 billion in the calendar year 2022.

[Tableau Dashboard – CLICK HERE](#)

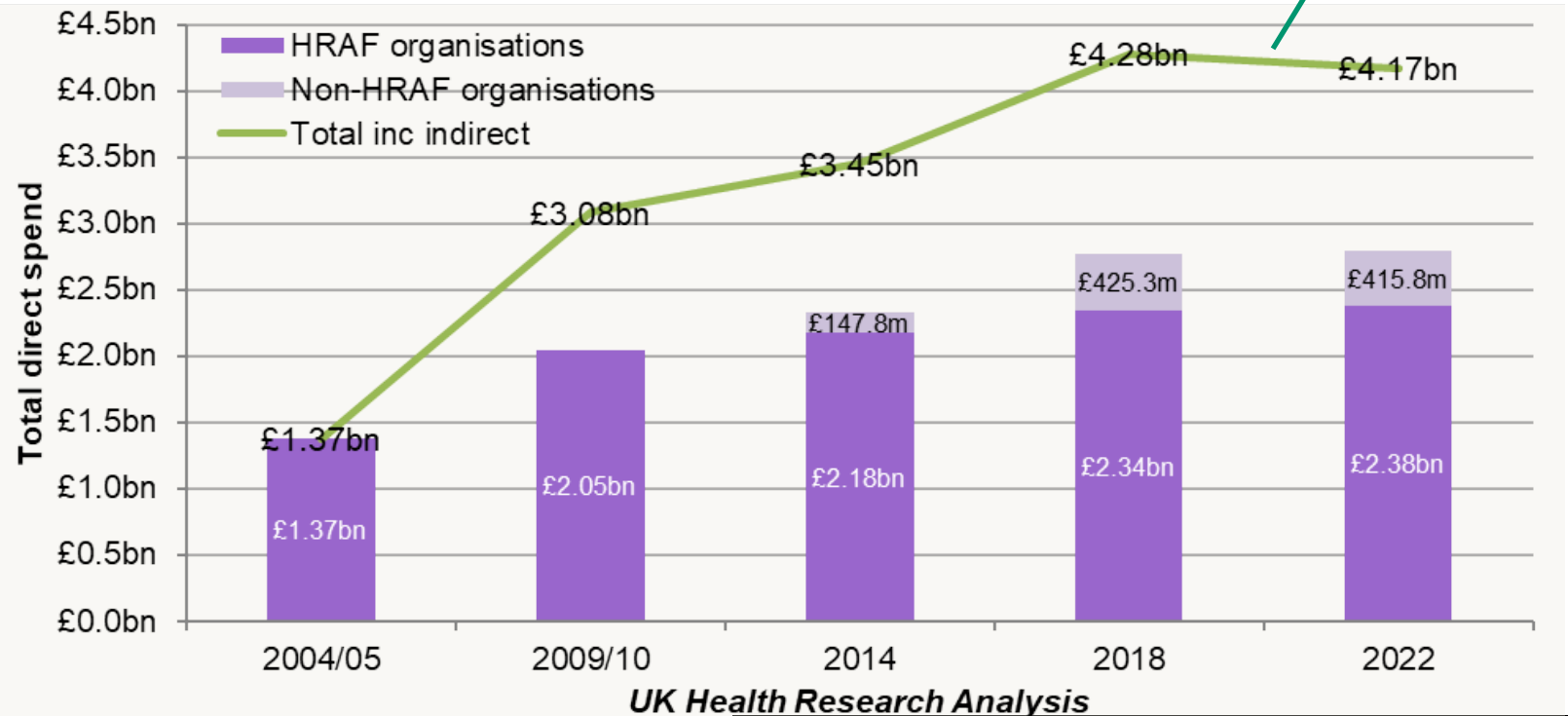
2022 analysis – overview (vs time)

- Comparisons across all five analyses in this series
- Participation has grown over time; from 12 to 173 organisations
- Original 12 funders (HRAF) accounts for ~90% spend
- Award value growth since 2004/05 has slowed

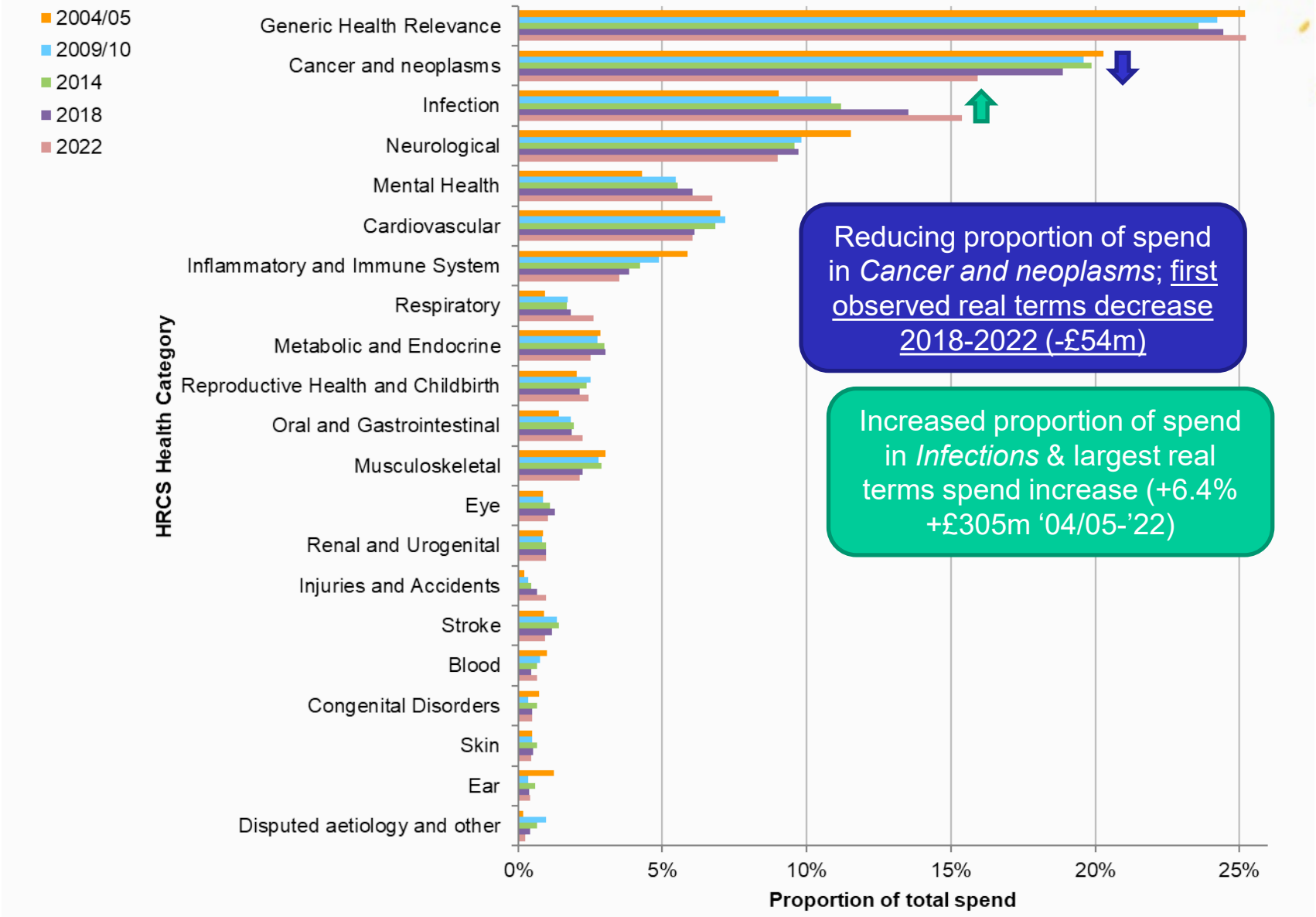
First observed decrease in real terms funding:

Direct: +£22m, 0.8%

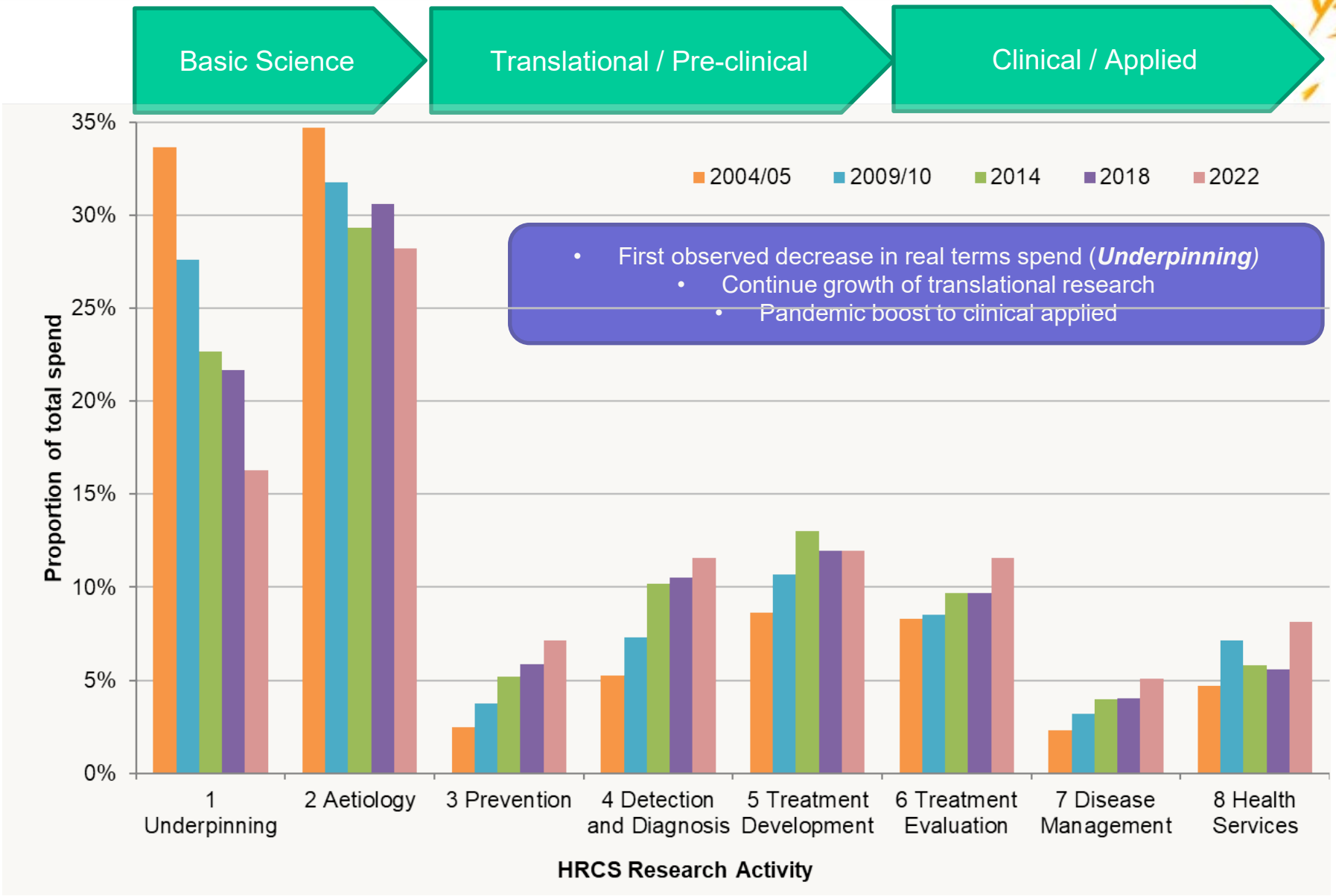
Indirect: -£133m, -9%



2022 analysis – Health Categories



2022 analysis – Research Activities



Impact of the analyses

- Evidence base for bespoke reporting
- Resource for development of automated coding (auto-coding) algorithms
- Referenced in a wide range of annual reports, publications and Government papers
- Increasing visibility of the ‘whole environment’ of funding
- Increased uptake of system internationally

Future of HRCS

- Improving the speed & frequency of analyses
- More bespoke ‘sub-analyses’ and data re-use
  - Prevention
  - UKRI health review
  - Place agenda
- Automated coding – WiP
  - Other developments to data automation

Igniting our potential

Health Research Classification System

Understanding the System

What is the HRCS?

Definition

"A system for classifying and analysing health and biomedical research funding"

Purpose

A common classification system applied to award funding which gives an overview of funding patterns covering the full spectrum of all types of research across all areas of health and disease

Uses

Gives meaningful comparisons across time, between different portfolios, **and** different organisations

Allows us to answer strategic questions about investment

Key aspects of HRCS coding process

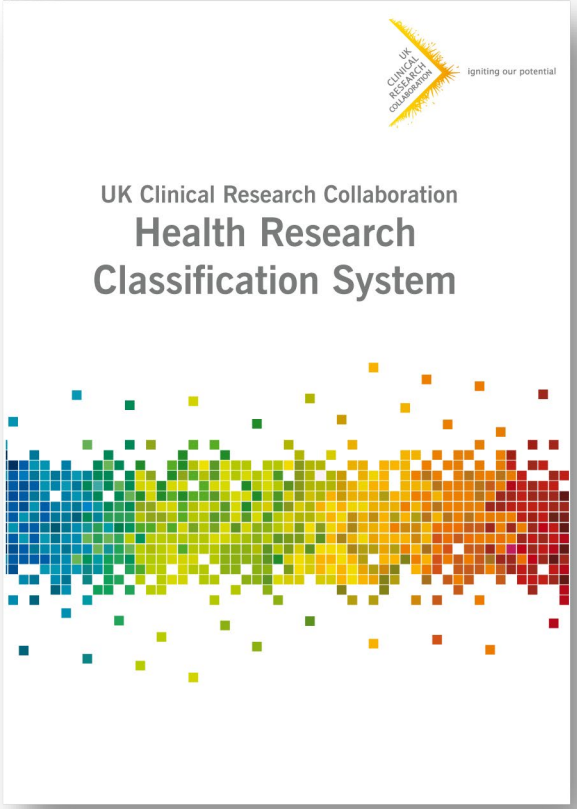
- **Tried and Tested**
  - Accumulated experience begun in 2005
  - Several major analyses, thousands of awards
  - Range of organisations, award types and settings

- **Stable and Consistent**
  - Approach suits all funders (low bar for entry)
  - HRCS is openly available (but not for alteration)
  - Simple reproducible rules (equal proportions, minimum number of codes, etc.)

Value lies in ability to compare over time and between settings

Structure of the HRCS

- Two dimensional system
  - Health Categories
  - Research Activity Codes
- Health Categories
  - All areas of health or disease
  - 21 individual categories
  - Based on **WHO International Classification of Diseases**
- Research Activity Codes
  - All types of research activity from basic to applied
  - 8 groups with 48 codes in X.X format
  - Based on cancer **Common Scientific Outline**



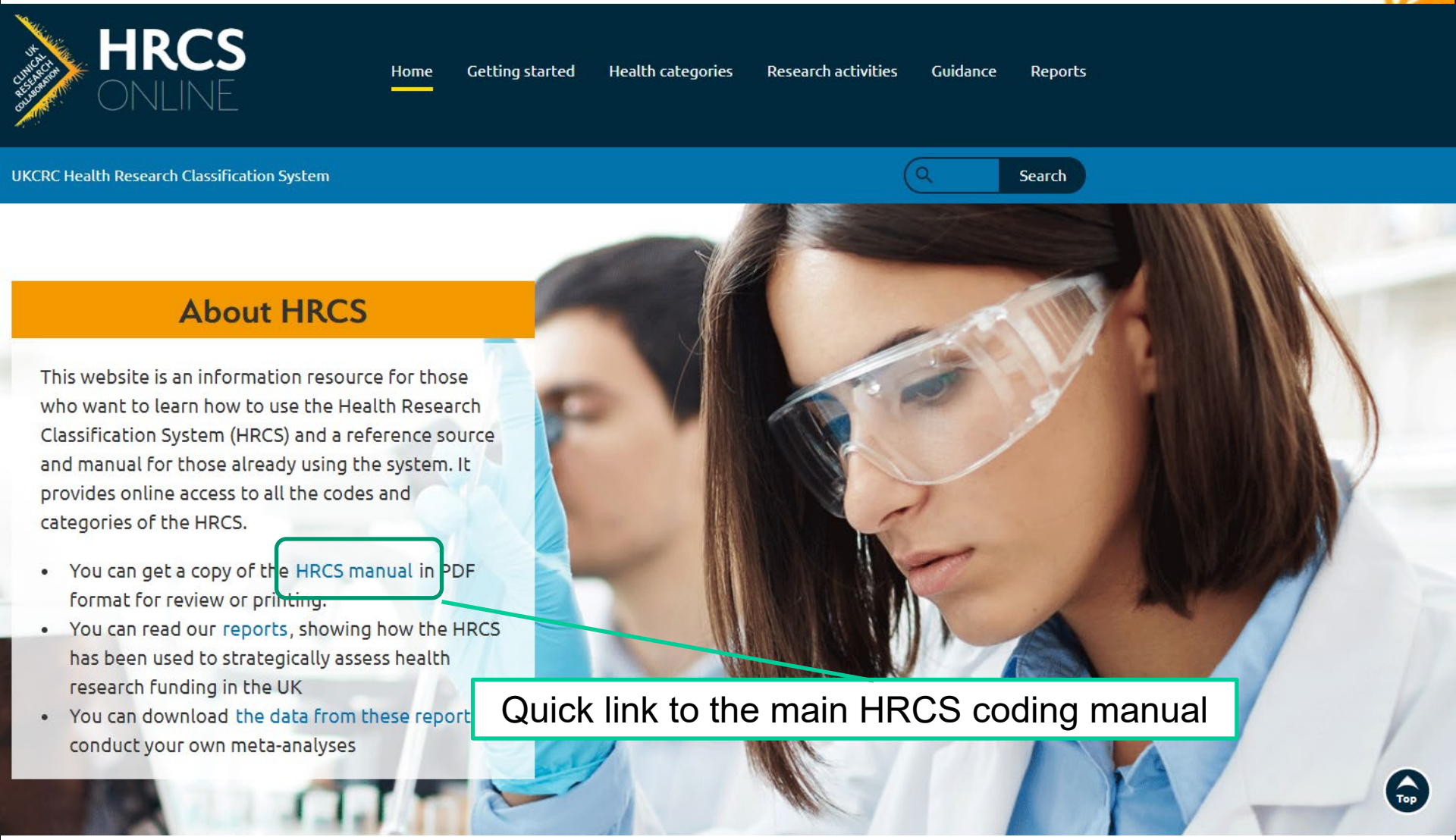
Key features of the HRCS

- Coding is based on the main research objective
  - Not an automatic keyword system
  - Does not capture all potential downstream outcomes
    - "within the lifetime of the award"
  - Can complement existing coding systems
  - Fit for purpose as overview tool (it can't do everything!)

- Coding is (usually) linked directly to associated funding
  - Codes based on lifetime of award & spend annualised to match
    - *Annualised Spend = (Commitment/Duration) * time active in analysis year*
  - Exact percentages with every code – equal spend weighting
    - Research management tool **not** a financial audit tool
    - FEC, supporting infrastructure, capital funding

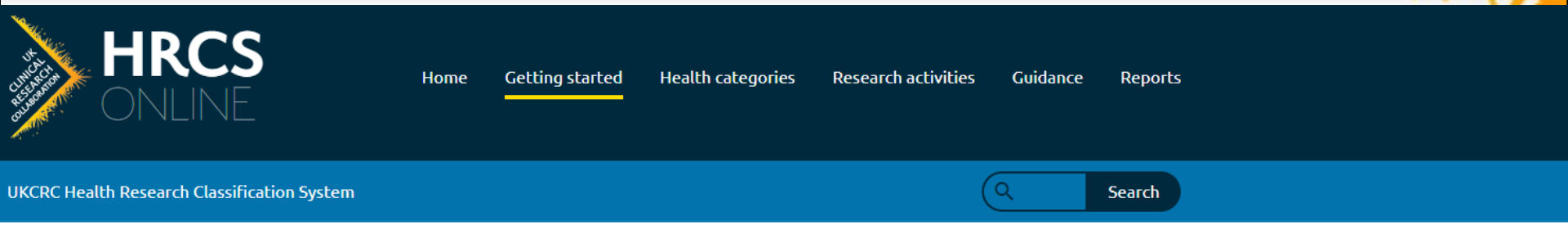
- System designed to provide broad overview of research
  - Provides a "one-size-fits" all approach to classification
  - Scientific expertise not required – technical distractions???

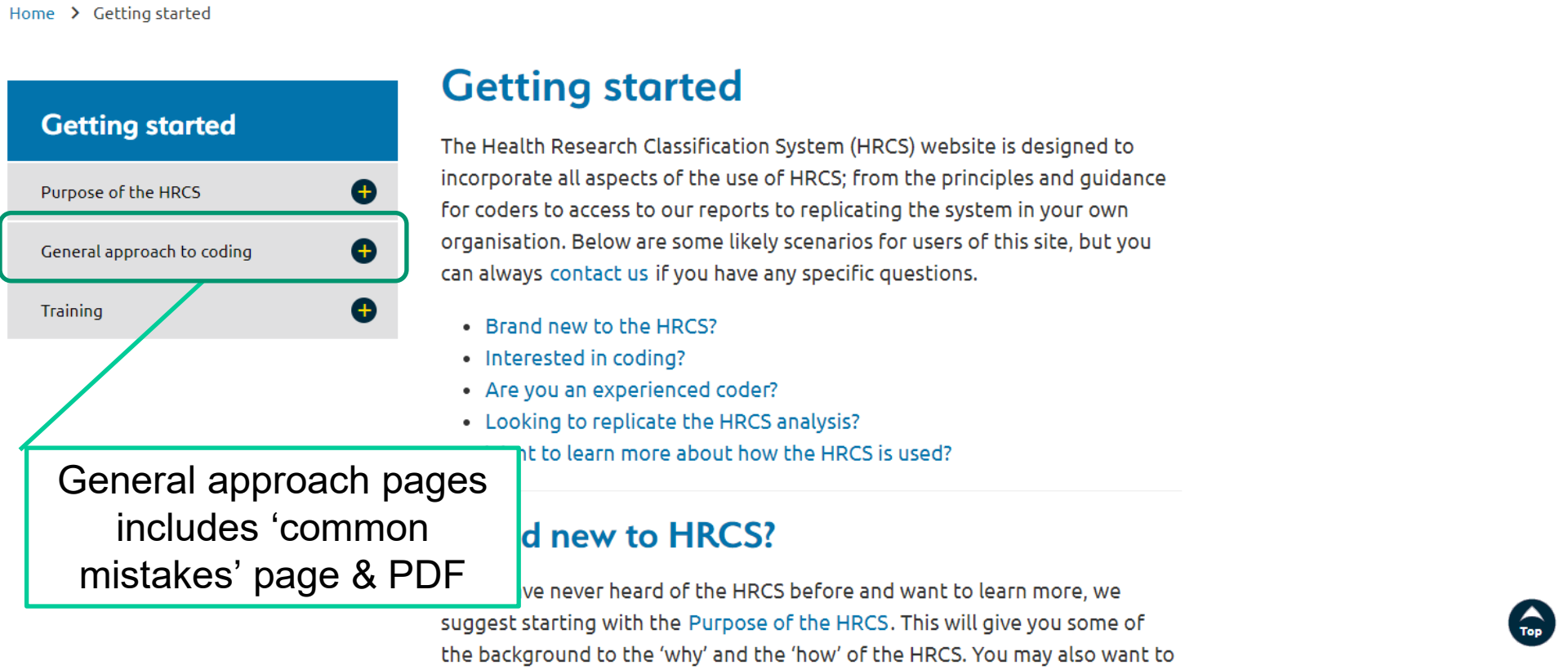
HRCS Wesbite - www.hrcsonline.net



Working in partnership. Changing cultures. Igniting our potential

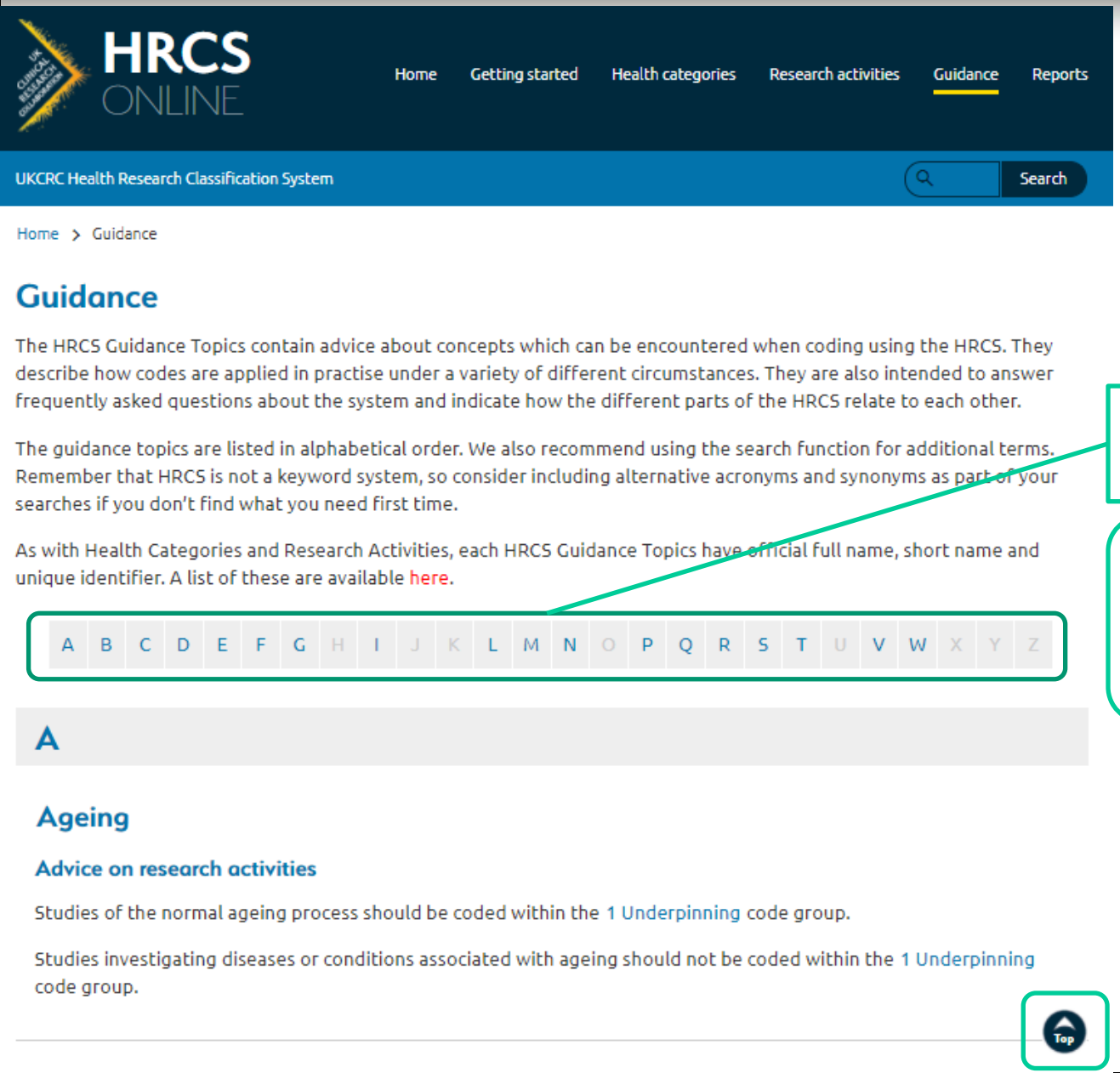
Getting Started





General approach pages includes 'common mistakes' page & PDF

Guidance Topics



Single page and A-to-Z quick-link bar

Top Tips:

- Use A-to-Z + 'Top' button
- Single page great for Ctrl+F
- or use new site search

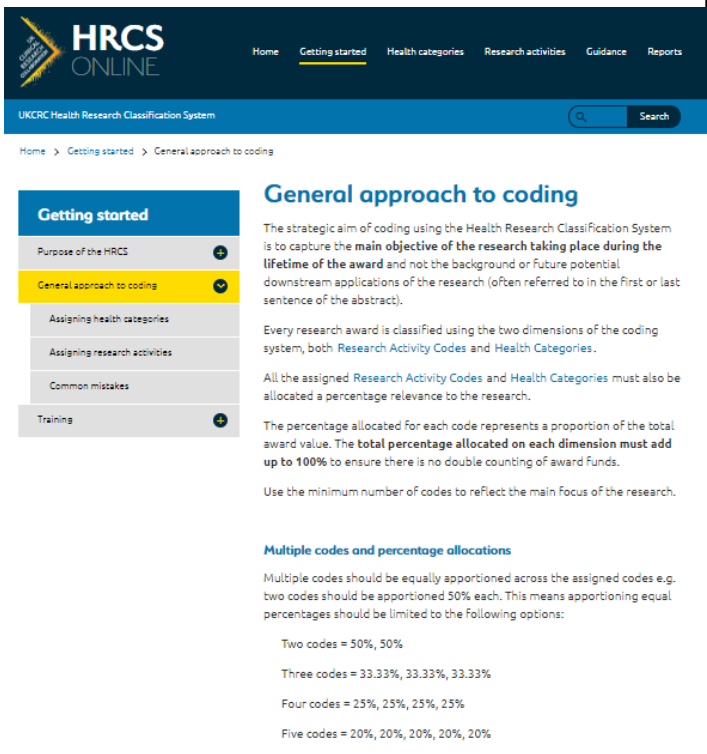
Igniting our potential

Health Research Classification System

How to code using HRCS

Approach to Coding – "How to code"

- Identify **main** aim(s) and health focus(es) of research **within lifetime of award**
  - Be wary of keywords or terms
- Allocate the **minimum** number of codes to reflect these
  - 1-5 Health Categories
  - 1-2 (max. 4) Research Activities
- Assign **equal** percentages adding to 100% for both Health Categories and Research Activity Codes
  - Analyses use this to assign equal spend weighting
- Remember HRCS context = landscaping tool
  - Used only to identify main research focus i.e. not auditing
  - Complements other systems for additional granularity



- Blood
- Cancer and neoplasms
- Cardiovascular
- Congenital Disorders
- Ear
- Eye
- Infection
- Inflammatory and Immune System
- Injuries and Accidents
- Mental Health
- Metabolic and Endocrine
- Musculoskeletal
- Neurological
- Oral and Gastrointestinal
- Renal and Urogenital
- Reproductive Health and Childbirth
- Respiratory
- Skin
- Stroke

- *Generic Health Relevance*
- *Disputed Aetiology and Other*

Each category includes normal/healthy **and** disease processes

Categories not always deducible from causation, symptoms, or location

Notes on Health Categories (1)

- Cancer and neoplasms
  - all types - not coded by site e.g. lung cancer
- Cardiovascular
  - includes atherosclerosis
- Congenital Disorders
  - multiple syndromes - excludes single focus syndromes like congenital heart disorders
- Infection
  - all types - not coded by site e.g. respiratory tract infections
- Inflammatory and Immune System
  - about immune system (not just immune response) - includes rheumatoid arthritis
- Musculoskeletal
  - includes osteoarthritis

Notes on Health Categories (2)

- Mental Health
  - includes normal behavioural and cognitive function and all abnormal conditions defined by behavior – distinct from *Neurological*
  - Depression, schizophrenia, psychosis and personality disorders, addiction, suicide, anxiety, eating disorders, learning disabilities, autistic spectrum disorders and studies of normal psychology, cognitive function and behavior
  - Any studies coded exclusively to *1.2 Psychological / Socioeconomic*
- Neurological
  - Brain function / 'wiring' – distinct from *Mental Health*
  - Includes neurodegenerative conditions: Dementias, Parkinson's, Alzheimer's.
  - Also includes TSEs, epilepsy, multiple sclerosis prions/BSE/CJD and studies of the normal brain and nervous system
  - Can include studies of circadian rhythm, headaches/migraines (usually, not exclusively)
  - Exclude studies of the brain with psychological conditions listed in *Mental Health*
- Examples of dual coding (N/MH)
  - Mapping of brain structure (should also be *1.1 Biological* and *1.2 Psychological*)

Notes on Health Categories (3)

- Reproductive Health and Childbirth
  - includes all aspects of pregnancy and the new born
- Respiratory
  - includes asthma

- Generic Health Relevance
  - **all** areas of health or general health
  - Or >5 Health Categories
- Disputed Aetiology and Other
  - A **few** very specific areas including Gulf War syndrome and chronic fatigue syndrome
  - Not a dustbin category

Guidance Topics on Health Categories

See specific guidance on:

- Sequelae
  - When a condition is a consequence or side effect of a pre-existing condition
  - If double coding, does one code make sense without the other?

- Multiple diseases and conditions
  - Health category list for use in four areas – Alcohol; Diet / nutrition; Physical activity / exercise; Tobacco / smoking
  - **Use only in the absence of other information**

1	**Underpinning Research**	**Research that underpins investigations** into the cause, development, detection, treatment and management of diseases, conditions and ill health	Discovery
2	**Aetiology**	Identification of determinants that are involved in the **cause, risk or development** of disease, conditions and ill health	
3	**Prevention of Disease and Conditions, and Promotion of Well-Being**	Research aimed at the primary **prevention** of disease, conditions or ill health, **or promotion of well-being**	Translational
4	**Detection, Screening and Diagnosis**	**Discovery, development and evaluation of diagnostic, prognostic and predictive markers and technologies**	
5	**Development of Treatments and Therapeutic Interventions**	Discovery and development of therapeutic interventions and testing **in model systems and preclinical settings**	
6	**Evaluation of Treatments and Therapeutic Interventions**	Testing and evaluation of therapeutic interventions **in clinical, community or applied settings**	Applied/Clinical
7	**Management of Diseases and Conditions**	Research into **individual** care needs and management of disease, conditions or ill health	
8	**Health and Social Care Services Research**	Research into the provision and delivery of health and social care services or health policy **at an organisational level**	

Notes on Research Activity Codes (1)

- **1 Underpinning**
  - normal / healthy / non-diseased
    - Excludes Cancer & Infection?
  - pain, immune responses, pregnancy, ageing, cell death DNA repair are considered normal
  - not just biological = chemistry, psychology, social
- **2 Aetiology**
  - “mechanism of disease”
  - not just causation – describing development, progression and life course of disease
  - includes epidemiology and observational studies

Notes on Research Activity Codes (2)

- **3 Prevention**
  - excludes secondary prevention (early detection) and tertiary prevention (prevention of a condition recurring)
- **4 Detection and Diagnosis**
  - markers / screening / monitoring / prediction
  - 4.1 – pre-clinical / lab based
  - 4.2 – clinical studies in humans
- **5 Treatment Development**
  - pre-clinical / lab based - including patient samples

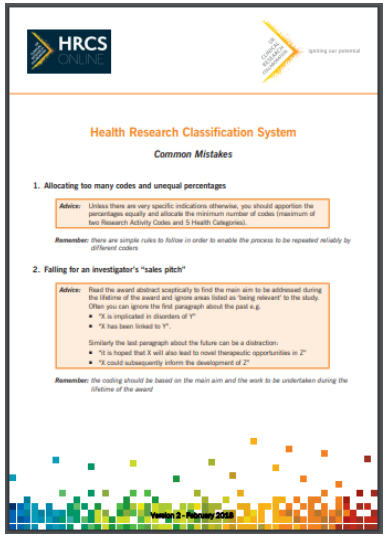
Notes on Research Activity Codes (3)

- **6 Treatment Evaluation**
  - clinical studies in humans – including all **therapeutic** trials phases I-IV
  - includes economic evaluation and assessing quality of life as part of study measures
- **7 Disease Management**
  - personal perspective – research into activities of health professionals and/or needs of patients
- **8 Health Services**
  - institutional perspective – research into organisations and service delivery

Guidance Topics on Research Activity Codes

See specific guidance on:

- **Repeated terms and concepts**
  - *e.g. trials, policy, education, evaluation etc.*
  - Methodology
    - *Appears in three groups (1,2,8)*
  - Resources and infrastructure
    - *Appears in all eight groups*


HRCS ONLINE

Health Research Classification System

Common Mistakes

1. Allocating too many codes and unequal percentages

Advice: Unless there are very specific indications otherwise, you should apportion the percentages equally and allocate the minimum number of codes (maximum of two Research Activity Codes and 5 Health Categories).

Remember: *there are simple rules to follow in order to enable the process to be repeated reliably by different coders*

2. Falling for an investigator's "sales pitch"

Advice: Read the award abstract sceptically to find the main aim to be addressed during the lifetime of the award and ignore areas listed as 'being relevant' to the study. Often you can ignore the first paragraph about the past e.g.

- "X is implicated in disorders of Y"
- "X has been linked to Y".

Similarly the last paragraph about the future can be a distraction:

- "It is hoped that X will also lead to novel therapeutic opportunities in Z"
- "X could subsequently inform the development of Z"

Remember: *the coding should be based on the main aim and the work to be undertaken during the lifetime of the award*

Version 2 - February 2018

- **Coding Mistakes document**
  - *Contains some common examples of coding errors*
  - *E.g. identifying preventative interventions (RA3) as treatments (RA5 or 6)*

Igniting our potential

Health Research Classification System

Coding Summary

Approach to Coding – Final Reminder

- Identify **main** aim(s) and health focus(es) of research **within lifetime of award**
- Allocate the **minimum** number of codes to reflect these
  - 1-5 Health Categories
    - Consider the 19 disease/physiological system codes before Generic/Other
  - 1-2 (max. 4) Research Activities
    - Chose the Research Activity Group before selecting a sub-code
- Assign **equal** percentages adding to 100% for both Health Categories and Research Activity Codes (unless there are clear reasons not to)

Rules For assigning Health Categories

- **EXAMPLE:** *A clinical trial in humans to test a new drug therapy for treating lung cancer*

- Identify the main health or disease focus(es) of the award
  - Ignore background work often listed as **being relevant** and future potential downstream outcomes
- Match each to a Health Category obeying the specific inclusion criteria
  - e.g. “testing a treatment for **lung cancer**” -> *Cancer* **not** *Respiratory*
  - Knowledge of pathogenesis, symptoms and disease site may not always be relevant
- Assign up to five Health Categories and apply equal percentages
  - *Generic Health Relevance* should be assigned if more than five categories apply or there is wide health relevance
  - The *Other* category is only for **very** specific cases – do not use if you are uncertain
- Be aware of special rules and guidance to help with certain topics
  - e.g. studies of alcohol, diet, exercise and smoking
  - e.g. consequences or side effects of a pre-existing condition

Rules for assigning Research Activity Codes

- **EXAMPLE (cont.):** *A clinical trial in humans to test a new drug therapy for treating lung cancer*

- Identify the main aim(s) of the award
  - Ignore background work often listed as **being relevant** and future potential downstream outcomes
- Match each to a Research Activity Code **group (1-8)**
  - e.g. “a trial in humans testing a new therapy” -> *6 Treatment Evaluation*
  - Note that HRCS code names and research concepts are repeated across code groups
    - e.g. **trials** are not always therapeutic trials
    - e.g. studies of **therapies** can be in humans or pre-clinical
- Then select appropriate sub-code from within that group
  - e.g. “it is a drug trial” -> *6.1 Pharmaceuticals*
- Assign up to two Research Activity Codes and equal percentages
  - There is an upper limit of four RACs for large programmes

Some final advice…

Don't panic!	It's a complex system, you won't be expected to learn it instantly
Be your own quality control	Re-check the awards you coded first once you're familiar with the system… do you agree with your past self?
Use a buddy system	Discuss tricky awards, potentially share your homework to get consistency with colleagues
Find experienced coders in your organisation	Your organisation may have some specific tips and tricks experienced coders are great sources for advice

Ask me… sparingly!

Igniting our potential

www.hrcsonline.net

james.carter@mrc.ukri.org