



Klassifisering av vitenskapelig publisering etter Health Research Classification System (HRCS)

En pilotstudie

Dag W. Aksnes

Arbeidsnotat 2016:12

Klassifisering av vitenskapelig publisering etter Health Research Classification System (HRCS)

En pilotstudie

Dag W. Aksnes

Arbeidsnotat 2016:12

Arbeidsnotat 2016:12

Utgitt av Nordisk institutt for studier av innovasjon, forskning og utdanning
Adresse Postboks 2815 Tøyen, 0608 Oslo. Besøksadresse: Økernveien 9, 0653 Oslo.

Prosjektnr. 12820418-26

Oppdragsgiver Norges forskningsråd
Adresse Postboks 564, 1327 Lysaker

Bilddesign Cathrine Årving
Foto Shutterstock

ISBN 978-82-327-0205-3
ISSN 1894-8200 (online)

www.nifu.no

Forord

På oppdrag fra Norges forskningsråd har NIFU gjennomført en pilotstudie for å utvikle en metodologi for å klassifisere norsk medisinsk og helsefaglig publisering etter det britiske Health Research Classification System (HRCS). Undersøkelsen er gjennomført innenfor en begrenset ressursramme, og det foreliggende arbeidsnotatet skisserer derfor kun i korte trekk et system for dette.

Oslo, 30.08.16

Sveinung Skule
Direktør

Susanne L. Sundnes
Forskningsleder

Innhold

1	Innledning	7
2	Beskrivelse av klassifiseringssystemet.....	10
2.1	Tidsskriftsbasert klassifisering	10
2.2	Klassifisering etter søkeord	14
3	Prosedyre	24
4	Testresultater	26
	Vedlegg – oversikt over HRCS-klassifisering av tidsskrifter.	29

1 Innledning

Problemstillingen om hvordan publikasjoner kan klassifiseres etter HRCS-systemet ble diskutert i NIFU arbeidsnotat 9/2014.¹ Som en innledning har vi nedenfor klippet ut relevant tekst fra dette notatet.

«I monitoreringssystemet er det et mål å kunne inndele forsknings- og innovasjonsindikatorer på en tematisk måte slik at de kan sammenholdes med helsedata som dekker sykdomsgrupper, helseutvikling og samfunnsbehov. Det britiske Health Research Classification System (HRCS) ivaretar dette behovet bedre enn disiplinindelinger og benyttes nå av Forskningsrådet og Kreftforeningen for å klassifisere søknader og prosjekter. De regionale helseforetakene har siden 2009 brukt systemet til å klassifisere prosjekter finansiert gjennom de regionale forskningsmidlene (både konkurranseutsatte og strategiske). I det nasjonale publiseringsystemet Cristin er det lagt inn en mulighet å registrere publikasjoner i henhold til HRCS-kategorier, men dette er ikke pålagt og i liten grad benyttet. I et pågående prosjekt ved Folkehelseinstituttet blir forskningsaktivitet uttrykt gjennom HRCS-koder sammenholdt med mål for dødelighet og sykkelighet. Dette arbeidet er en del av det norske sykdomsbyrdeprosjektet ved Nasjonalt folkehelseinstitutt.

HRCS er ikke begrenset til fagområdet medisin, men inkluderer all forskning knyttet til temaet helse innenfor alle fag og disipliner. Det er to dimensjoner i HRCS – forskningstype og helsekategorier, der det kun er sistnevnte dimensjon som omtales i dette notatet, primært fordi det er helsekategoriene som er mest relevante og minst vanskelige å innføre for ressurs- og resultatindikatorer. Helsekategoriene i den engelskspråklige versjonen er:

1. Blood
2. Cancer
3. Cardiovascular
4. Congenital Disorders
5. Ear
6. Eye
7. Infection
8. Inflammatory and Immune System
9. Injuries and Accidents
10. Mental Health
11. Metabolic and Endocrine
12. Musculoskeletal
13. Neurological
14. Oral and Gastrointestinal
15. Renal and Urogenital
16. Reproductive Health and Childbirth
17. Respiratory
18. Skin
19. Stroke
20. Generic Health Relevance

¹ Gunnar Sivertsen, Fredrik Piro, Markus M. Bugge, Arne M. Fevolden, Espen Solberg, Ole Wiig: *Monitoreringssystem for forskning og innovasjon innen helse og omsorg. Et forprosjekt for å undersøke mulighetene.*

21. Other

Det er i utgangspunktet ikke enkelt å bruke HRCS på forsknings- og innovasjonsindikatorer, det er enklere i sammenheng med behandling av søknader og prosjekter. HRCS er ikke like meningsfull å bruke innenfor pleie og omsorg som innenfor sykehussektoren ettersom primærhelsetjenesten har andre overordnede målsettinger enn spesialisthelsetjenesten. Flere av respondentene i NIFUs arbeidsnotat viser til forskning og områder som ikke (uten videre) kan kategoriseres i henhold til HRCS. Eksempler på slike områder er forskning på ulike helsetjenester og samfunnsøkonomiske analyser.

HRCS for bibliometriske indikatorer

Per i dag er det ikke mulig å analysere verken publisering eller hvor mye publikasjonene er sitert i henhold til sykdomskategoriene i HRCS. Analyser som ser på publisering eller siteringer innen og på tvers av fag har til nå fulgt Thomson Reuters faginndeling, som er mer omfattende enn HRCS og som heller ikke lar seg omkode til HRCS, fordi den i utgangspunktet er disiplinorientert. I Universitets- og høgskolerådets (UHR) faginndeling av tidsskrifter som ligger i registeret for Cristin hos NSD, foreligger en enklere inndeling som også er disiplinbasert. Denne dekker langt flere tidsskrifter enn hos Thomson Reuters.

Vi vurderer det slik at det er mulig å lage en ny inndeling basert på HRCS av relevante tidsskrifter i begge de nevnte registrene. Det vil da være mulig å bruke systemet til bibliometriske undersøkelser som omfatter artikler i vitenskapelige tidsskrifter. Innen helse og omsorg er dette den vanligste vitenskapelige publikasjonstypen. Reklassifisering av tidsskrifter forutsetter et utviklings- og utprøvningsarbeid. Deretter vil klassifiseringen enkelt kunne oppdateres fra år til år. Bibliometriens begrensninger vil likevel fortsatt gjelde: Det er primært FoU-aktivitet i offentlig sektor som blir offentliggjort i vitenskapelige tidsskrifter. Næringslivet publiserer mye sjeldnere, men på den annen side er bedrifter med helserettet FoU-aktivitet relativt mer aktive i publisering enn andre FoU-aktive bedrifter.»

Som beskrevet ovenfor, er det behov for et utviklingsarbeid for å lage et system for klassifisering av publikasjonene etter HRCS-kategorier. Så langt vi kjenner til, foreligger det ikke etablert metodologi for dette internasjonalt. Analysene som har vært gjennomført i Storbritannia synes å ha tatt utgangspunkt i finansieringskilder, og publikasjoner er så koblet til HRCS-kategorier på bakgrunn av hvem som har finansiert forskningen. Det foreligger derimot ikke noe system for direkte klassifisering av publikasjoner basert på tidsskrift eller tittel. Den vanlige metoden for klassifisering av publikasjoner innenfor bibliometri er basert på tidsskriftsklassifisering, slik at alle publikasjoner i et bestemt tidsskrift blir klassifisert i ett eller flere fagfelt. Thomson Reuters/Web of Science opererer med et klassifiseringssystem på om lag 250 fagfelt. Noen av disse, men langt fra alle, vil samsvare godt med HRCS-systemet. For eksempel er det en egen kategori for onkologi som da vil kunne klassifiseres under HRCS-cancer.

I notatet har vi utviklet et klassifiseringssystem som dels baseres på tidsskriftsnavn og etablerte WoS-kategorier, dels på bruk av søkeord/nøkkelord. Det er et betydelig antall tidsskrifter innen medisin og helse som dekker tema som ikke direkte lar seg klassifisere innen HRCS-systemet. Eksempler er generelle medisinske tidsskrifter slik som *Lancet* og generelle biomedisinske tidsskrifter slik som *Cell*. Videre er det mange tidsskrifter som publiserer forskning innen flere enn én HRCS kategori eller som har et tema som ikke direkte passer inn i HRCS-klassifiseringssystemet. Eksempler på slike tidsskrifter er *BMC Pediatrics* og *Anesthesiology*.

Klassifiseringssystemet er beskrevet nærmere i neste kapittel. Systemet er utviklet for å kunne anvendes mekanisk, hvor det ikke benyttes fagekspertise i klassifiseringen av enkeltpublikasjoner. Her ligger det også begrensninger. Testing av systemet viser at det er vanskelig å unngå at en del enkeltpublikasjoner blir feilklassifisert. Systemet er imidlertid utviklet for å kunne anvendes på makronivå, dvs. analyser av nasjonale publiseringsmønstre. Da blir slike feilkilder av mindre betydning. Det bør presiseres at feilkildene gjør at det er mindre egnet til å analysere små forskningsenheter med få publikasjoner.

Et annet problem er at det ikke foreligger detaljert informasjon om hvordan de ulike HRCS-kategoriene er avgrenset. En del informasjon er listet på web-sidene til HRCS (<http://www.hrcsonline.net/hc>), men

denne er ikke tilstrekkelig til å utarbeide et komplett sett med søkeord. Derfor involverer prosessen med å utvikle systemet en ikke ubetydelig grad av skjønn.

Prosjektet har tatt utgangspunkt i tidsskrifter og publikasjoner som er indeksert i Web of Science. Denne databasen dekker en stor majoritet av den vitenskapelige publiseringen innenfor medisin og helsefeltet, men blant annet inngår ikke norskspråklige tidsskrifter. Systemet kan eventuelt videreutvikles til å omfatte også øvrig publisering. Da er det imidlertid behov for å tilpasse søkesystemet til norskspråklig terminologi, og dette inngår ikke i prosjektet.

Systemet er utviklet som en pilotstudie, som det vil være mulig å justere og raffinere basert på erfaringer og innspill. Før systemet eventuelt implementeres, foreslås det at relevante aktører gis mulighet til å komme med synspunkter og vurderinger, dette som en kvalitetssikring av systemet. Notatet inneholder også noen forslag til hvordan systemet kan videreutvikles. Dette er også elementer som bør vurderes i denne sammenheng.

2 Beskrivelse av klassifiseringssystemet

Å definere eller avgrense et fagfelt kan være utfordrende og problematisk – det er et generelt problem som ikke bare oppstår innen bibliometri. Selv blant fagekspertene kan det noen ganger være stor uenighet om hvordan et bestemt fagfelt best kan defineres. En økende grad av tverrfaglig forskning gjør det også stadig vanskeligere å trekke grensene mellom ulike fagfelt. I bibliometri oppstår spesielle problemer fordi mange vitenskapelige artikler dreier seg om komplekse tema som ikke entydig kan klassifiseres innen en bestemt disiplin eller fagfelt. For eksempel kan en og samme artikkel legitimt kunne klassifiseres som for eksempel biokjemi, molekylærbiologi, mikrobiologi, immunologi og klinisk medisin. På grunn av dette vil artikler kunne klassifiseres innen flere enn én HRCS-kategori. I klassifiseringssystemet som er utviklet, er det også lagt opp til dette. For eksempel vil artikkelen, “Dental anxiety and dental attendance among 25-year-olds in Norway: time trends from 1997 to 2007”, både klassifiseres innen 10. Mental Health og innen 14. Oral and Gastrointestinal (som også omfatter odontologi). Dette er i tråd med retningslinjene for HRCS-systemet, men det presiseres samtidig: “Use the minimum number of codes to reflect the main focus of the research”.

Generelt vil ofte noen artikler være enkle å klassifisere, de befinner seg innenfor kjernen av et fagfelt, mens andre er vanskeligere, de er bare svakt relatert til et fagfelt eller befinner seg i grenselandet mellom flere forskjellige fagfelt. Ideelt sett bør kriteriene anvendt inkludere alle publikasjoner som hører til et bestemt fagfelt, og samtidig være presise nok til å utelukke publikasjoner irrelevante for feltet. De fleste bibliometriske analyser benytter tidsskriftsbaserte definisjoner, noe som betyr at alle artikler i et gitt tidsskrift blir tilordnet det samme fagfeltet. Et problem med denne metoden er at mange tidsskrifter inneholder artikler som omhandler flere ulike tema. Følgelig vil noen artikler bli inkludert i et fagfelt som de bare er svakt relatert til, mens andre, mer relevante artikler ikke kommer med. På bakgrunn av dette er det i systemet lagt opp til en kombinasjon av tidsskrifts- og søkeordbasert metodologi.

2.1 Tidsskriftsbasert klassifisering

I Web of Science (WoS) er tidsskriftene klassifisert i om lag 250 forskjellige fagfelt dekkende hele spekteret av disipliner fra humaniora til naturvitenskap og teknologi. Til HRCS-klassifiseringssystemet vurderes det som tilstrekkelig å inkludere tidsskriftene som er klassifisert innen klinisk medisin, biomedisin, psykologi og helsefag. I tillegg er det relevant å ta med tidsskriftene som er klassifisert innen ernæring (Nutrition & dietetics), som er klassifisert i den overordnede kategorien Agriculture and food science. Videre er det relevant å ta med generelle vitenskapelige tidsskrifter som publiserer forskning innen en rekke fagfelt slik som *Plos One*. Tabell 1 viser hvilke WoS-kategorier som er relevante å inkludere i målesystemet.

Tabell 1. WoS-kategorier relevante å inkludere i målesystemet

WoS-disiplin	WoS-underdisiplin	WoS-disiplin	WoS-underdisiplin	
CLINICAL MEDICINE	ALLERGY	BASIC LIFE SCIENCES	BIOCHEMICAL RESEARCH METHODS	
	ANDROLOGY		BIOCHEMISTRY & MOLECULAR BIOLOGY	
	ANESTHESIOLOGY		BIOPHYSICS	
	AUDIOLOGY & SPEECH-LANGUAGE PATHOLOGY		BIOTECHNOLOGY & APPLIED MICROBIOLOGY	
	CARDIAC & CARDIOVASCULAR SYSTEMS		CELL & TISSUE ENGINEERING	
	CLINICAL NEUROLOGY		CELL BIOLOGY	
	CRITICAL CARE MEDICINE		DEVELOPMENTAL BIOLOGY	
	DENTISTRY/ORAL SURGERY & MEDICINE		GENETICS & HEREDITY	
	DERMATOLOGY		MICROBIOLOGY	
	EMERGENCY MEDICINE		REPRODUCTIVE BIOLOGY	
	ENDOCRINOLOGY & METABOLISM		CHEMISTRY, MEDICINAL	
	GASTROENTEROLOGY & HEPATOLOGY		BASIC MEDICAL SCIENCES	ENGINEERING, BIOMEDICAL
	HEMATOLOGY			MATERIALS SCIENCE, BIOMATERIALS
	INFECTIOUS DISEASES			MEDICAL INFORMATICS
	MEDICINE, GENERAL & INTERNAL	BIOLOGICAL SCIENCES	BEHAVIORAL SCIENCES	
	OBSTETRICS & GYNECOLOGY	BIOMEDICAL SCIENCES	ANATOMY & MORPHOLOGY	
	ONCOLOGY		IMMUNOLOGY	
	OPHTHALMOLOGY		INTEGRATIVE & COMPLEMENTARY MEDICINE	
	ORTHOPEDICS		MEDICAL LABORATORY TECHNOLOGY	
	OTORHINOLARYNGOLOGY		MEDICINE, RESEARCH & EXPERIMENTAL	
	PARASITOLOGY		NEUROIMAGING	
	PATHOLOGY		NEUROSCIENCES	
	PEDIATRICS		PHARMACOLOGY & PHARMACY	
	PERIPHERAL VASCULAR DISEASE		PHYSIOLOGY	
	PRIMARY HEALTH CARE		RADIOLOGY, NUCLEAR MEDICINE & MEDICAL IMAGING	
	PSYCHIATRY		TOXICOLOGY	
	PUBLIC, ENVIRONMENTAL & OCCUPATIONAL HEALTH		VIROLOGY	
	RESPIRATORY SYSTEM		PSYCHOLOGY	PSYCHOLOGY, APPLIED
	RHEUMATOLOGY			PSYCHOLOGY, BIOLOGICAL
	SURGERY			PSYCHOLOGY, CLINICAL
	TRANSPLANTATION			PSYCHOLOGY, DEVELOPMENTAL
	TROPICAL MEDICINE	PSYCHOLOGY, EXPERIMENTAL		
	UROLOGY & NEPHROLOGY	PSYCHOLOGY, MATHEMATICAL		
GERIATRICS & GERONTOLOGY	PSYCHOLOGY, MULTIDISCIPLINARY			
GERONTOLOGY	PSYCHOLOGY, PSYCHOANALYSIS			
HEALTH CARE SCIENCES & SERVICES	PSYCHOLOGY, SOCIAL			
HEALTH POLICY & SERVICES	AGRICULTURE & FOOD SCI	NUTRITION & DIETETICS		
NURSING	MULTIDISCIPLINARY JOURNALS	MULTIDISCIPLINARY SCIENCES		
REHABILITATION				
SPORT SCIENCES				
SUBSTANCE ABUSE				

Noen av WoS-kategoriene dekker imidlertid også tidsskrifter innen biologi og ikke-humanmedisin. Disse tidsskriftene bør ekskluderes. Som et vedlegg til rapporten finnes en excel-fil med en oversikt over alle tidsskrifter som inngår i de ulike fagfeltene. Her er det ekskludert en del tidsskrifter som faller utenfor humanmedisin. Tabell 2 viser antall inkluderte tidsskrifter per WoS-kategori. Som det fremgår, varierer kategoriene mye i størrelse i form av antall tidsskrifter inkludert.

Tabell 2. Antall inkluderte tidsskrifter per WoS-kategori.

WoS-kategori	Antall tidsskrift	WoS-kategori	Antall tidsskrift
BIOCHEMISTRY & MOLECULAR BIOLOGY	443	PERIPHERAL VASCULAR DISEASE	82
PHARMACOLOGY & PHARMACY	405	RESPIRATORY SYSTEM	79
NEUROSCIENCES	350	PSYCHOLOGY, DEVELOPMENTAL	79
ONCOLOGY	286	PSYCHOLOGY, SOCIAL	76
PSYCHOLOGY, MULTIDISCIPLINARY	276	GERIATRICS & GERONTOLOGY	65
PSYCHIATRY	271	OTORHINOLARYNGOLOGY	63
CELL BIOLOGY	255	CHEMISTRY, MEDICINAL	61
CLINICAL NEUROLOGY	241	SUBSTANCE ABUSE	56
IMMUNOLOGY	241	MEDICAL LABORATORY TECHNOLOGY	53
GENETICS & HEREDITY	210	MEDICINE, GENERAL & INTERNAL	50
BIOTECHNOLOGY & APPLIED MICROBIOLOGY	203	SURGERY	49
CARDIAC & CARDIOVASCULAR SYSTEMS	183	VIROLOGY	46
RADIOLOGY, NUCLEAR MEDICINE & MEDICAL IMAGING	182	DEVELOPMENTAL BIOLOGY	46
MEDICINE, RESEARCH & EXPERIMENTAL	179	PARASITOLOGY	45
ENDOCRINOLOGY & METABOLISM	172	MATERIALS SCIENCE, BIOMATERIALS	43
MICROBIOLOGY	157	ANESTHESIOLOGY	43
REHABILITATION	149	BEHAVIORAL SCIENCES	41
PSYCHOLOGY, CLINICAL	142	ALLERGY	40
NURSING	135	ANATOMY & MORPHOLOGY	38
TOXICOLOGY	123	GERONTOLOGY	33
OBSTETRICS & GYNECOLOGY	122	MEDICAL INFORMATICS	33
ENGINEERING, BIOMEDICAL	120	REPRODUCTIVE BIOLOGY	33
PHYSIOLOGY	119	CRITICAL CARE MEDICINE	31
PUBLIC, ENVIRONMENTAL & OCCUPATIONAL HEALTH	118	EMERGENCY MEDICINE	30
HEMATOLOGY	115	INTEGRATIVE & COMPLEMENTARY MEDICINE	29
DENTISTRY/ORAL SURGERY & MEDICINE	114	AUDIOLOGY & SPEECH-LANGUAGE PATHOLOGY	26
NUTRITION & DIETETICS	113	PEDIATRICS	25
BIOPHYSICS	110	CELL & TISSUE ENGINEERING	24
UROLOGY & NEPHROLOGY	109	RHEUMATOLOGY	23
GASTROENTEROLOGY & HEPATOLOGY	109	PSYCHOLOGY, PSYCHOANALYSIS	20
SPORT SCIENCES	104	NEUROIMAGING	19
HEALTH CARE SCIENCES & SERVICES	99	MULTIDISCIPLINARY SCIENCES	17
ORTHOPEDICS	96	PATHOLOGY	17
BIOCHEMICAL RESEARCH METHODS	94	PSYCHOLOGY, BIOLOGICAL	16
INFECTIOUS DISEASES	94	PSYCHOLOGY, MATHEMATICAL	14
PSYCHOLOGY, APPLIED	91	ANDROLOGY	8
PSYCHOLOGY, EXPERIMENTAL	91	MICROSCOPY	6
HEALTH POLICY & SERVICES	88	PRIMARY HEALTH CARE	5
OPHTHALMOLOGY	88	TRANSPLANTATION	5
DERMATOLOGY	87		

Tabell 3 viser hvordan WoS-kategoriene er klassifisert etter HRCS-systemet. Det er imidlertid mange WoS-kategorier innen medisin og helsefag som ikke lar seg klassifisere etter HRCS-systemet, disse er ikke listet i tabellen. Typisk dreier dette seg om fagfelt som dekker flere enn én HRCS-kategori (f.eks. Physiology) eller mer basale biomedisinske kategorier (f.eks. Cell Biology).

Tabell 3. Fordeling av WoS-kategorier til HRCS-kategorier.*

WoS-kategori	HRCS-kategori																				
	1. Blood	2. Cancer	3. Cardiovascular	4. Congenital Disorders	5. Ear	6. Eye	7. Infection	8. Inflammatory and Immune System	9. Injuries and Accidents	10. Mental Health	11. Metabolic and Endocrine	12. Musculoskeletal	13. Neurological	14. Oral and Gastrointestinal	15. Renal and Urogenital	16. Reproductive Health and Childbirth	17. Respiratory	18. Skin	19. Stroke	20. Generic Health Relevance	21. Other
ALLERGY							x														
ANDROLOGY														x							
AUDIOLOGY & SPEECH-LANGUAGE PATHOLOGY				x																	
CARDIAC & CARDIOVASCULAR SYSTEMS		x																			
CLINICAL NEUROLOGY												x									
DENTISTRY/ORAL SURGERY & MEDICINE													x								
DERMATOLOGY																	x				
ENDOCRINOLOGY & METABOLISM										x											
INFECTIOUS DISEASES						x															
OBSTETRICS & GYNECOLOGY															x						
ONCOLOGY	x																				
OPHTHALMOLOGY					x																
ORTHOPEDICS											x										
PERIPHERAL VASCULAR DISEASE		x																			
PSYCHIATRY									x												
RESPIRATORY SYSTEM																	x				
UROLOGY & NEPHROLOGY														x							
GASTROENTEROLOGY & HEPATOLOGY													x								
HEMATOLOGY	x																				
OTORHINOLARYNGOLOGY				x									x								
PARASITOLOGY						x															
SUBSTANCE ABUSE									x												
IMMUNOLOGY							x														
REPRODUCTIVE BIOLOGY															x						
NEUROIMAGING												x									
NEUROSCIENCES												x									
HEALTH POLICY & SERVICES																				x	
BEHAVIORAL SCIENCES									x												
INTEGRATIVE & COMPLEMENTARY MEDICINE																				x	
MEDICAL INFORMATICS																					x

*) I noen tilfeller bør tidsskriftene ikke overføres mekanisk. Følgende justeringer bør gjøres:

- HEMATOLOGY: Tidsskrifter som spesifikt omhandler leukemi/blodkreft klassifiseres som 2. Cancer, hvite blodceller som 8. Inflammatory and Immune System.
- OTORHINOLARYNGOLOGY**. Rene otologiske tidsskrifter klassifiseres under 5. Ear. Øvrige under 14. Oral and Gastrointestinal.

- IMMUNOLOGY: Tidsskrifter for astmaforskning klassifiseres under 17. Respiratory.
- CARDIAC & CARDIOVASCULAR SYSTEMS. Tidsskrifter omhandlende slag, klassifiseres under 19. Stroke.
- GERIATRICS & GERONTOLOGY. Klassifiseres etter tema, generelle under 20. Generic Health Relevance.
- NURSING: Klassifiseres etter tema, generelle under 20. Generic Health Relevance.
- SUBSTANCE ABUSE: I henhold til veiledningen til HRCS-systemet, skal alkoholforskning fordeles på fire ulike HRCS-kategorier. WoS-kategorien «Substance Abuse» dekker imidlertid primært sosiale og psykologiske problemer knyttet til rus. Derfor er kategorien plassert under 10. Mental Health.
- Tidsskrifter omhandlende injury/trauma (men ikke psykologisk) klassifiseres under 9. Injuries and Accidents.

Nærmere informasjon om hva som dekkes av de ulike WoS-kategoriene kan finnes på følgende website: http://ip-science.thomsonreuters.com/mjl/scope/scope_scie/

Resultatet av denne klassifiseringen for enkelttidsskrifter er vist i en fil som utgjør et vedlegg til rapporten. Tabell 4 viser antall tidsskrifter per HRCS-kategori basert på klassifiseringen ovenfor. Det er betydelig variasjon i antall tidsskrifter per kategori, og for en av kategoriene (4. Congenital Disorders) ble det ikke identifisert relevante spesialtidsskrifter.

Tabell 4. Antall inkluderte tidsskrifter per HRCS-kategori.

HRCS-kategori	Antall tidsskrifter	HRCS-kategori	Antall tidsskrifter
1. Blood	97	12. Musculoskeletal	137
2. Cancer	289	13. Neurological	516
3. Cardiovascular	243	14. Oral and Gastrointestinal	264
4. Congenital Disorders	0	15. Renal and Urogenital	117
5. Ear	36	16. Reproductive Health and Childbirth	150
6. Eye	88	17. Respiratory	82
7. Infection	149	18. Skin	80
8. Inflammatory and Immune Systems	212	19. Stroke	8
9. Injuries and Accidents	38	20. Generic Health Relevance	552
10. Mental Health	1059	Uklassifisert	2526
11. Metabolic and Endocrine	168		

2.2 Klassifisering etter søkeord

I tillegg til den tidsskriftsbaserte avgrensingen, hvor klassifiseringen skjer etter tidsskrift og ikke artikkelinnhold, involverer systemet bruk av søkeord. Det søkes i tittelen og sammendraget (abstractet) til alle publikasjoner, både artikler i tidsskrifter som allerede er klassifisert etter HRCS-kategorier og i tidsskrifter som ikke er det. Metoden med bruk av nøkkelord er anvendt av NIFU gjennom flere bibliometriske studier i andre områder, blant annet polarforskning og fiskeriforskning. Det finnes også mange eksempler på bruk av metoden internasjonalt, for eksempel en analyse av Welcome Trust finansiert forskning.²

For hver HRCS-kategori er det identifisert et sett med søkeord som skal brukes i denne prosessen. I prinsippet kan det for hver HRCS-kategori lages et stort antall søkeord. Vi har likevel begrenset antallet for hver kategori til ord og begreper som er sentrale og som det kan antas er nevnt i tittelen

² Jon Sussex, Yan Feng, Jorge Mestre-Ferrandiz, Michele Pistollato, Marco Hafner, Peter Burridge and Jonathan Grant (2016). Quantifying the economic impact of government and charity funding of medical research on private research and development funding in the United Kingdom. *BMC Medicine*, 14:32.

eller sammendraget til artiklene. Søkeordene bør identifisere artiklene på en entydig måte, det vil si at ord som ikke er spesifikke for én HRCS-kategori ideelt bør unngås. Videre bør ord som kan ha flere betydninger ikke brukes. For HRCS-kategori, 6. Eye, vil ordet «lens» for eksempel være problematisk å bruke siden dette også vil identifisere artikler hvor mikrosoplinsler og andre linser er omtalt. I valget av søkeord er det derfor ikke tilstrekkelig å identifisere sentrale begreper innen et felt, ordenes egnethet som søkeord må også vurderes.

For å identifisere søkeord innen de ulike HRCS-kategoriene, er det gjort en ordfrekvensanalyse av publikasjonstitler/abstract som allerede er klassifisert innen et bestemt HRCS-område basert på tidsskriftsnavn. Videre er tema som er listet som eksempler i HRCS-veiledningen brukt. I enkelte tilfeller ble også International Statistical Classification of Diseases and Related Health Problems (<http://www.who.int/classifications/icd/en/>) og Wikipedia benyttet. Valget av ord kan synes litt tilfeldig, noen sykdommer er nevnt ved navn, andre ikke. Dette kan for eksempel skyldes at sykdommene er spesifisert i HRCS-veiledningen eller at de kommer høyt opp i ordfrekvensanalysen. Motivet har altså ikke vært å identifisere et komplett sett med begreper, kun et sett med begreper som er egnet til å identifisere de aktuelle feltene bibliometrisk. Noen felt er enklere å avgrense enn andre, derfor varierer antallet ord mye fra kategori til kategori.

Nedenfor finnes en oversikt over søkeord for hver av de ulike HRCS-kategoriene. Innledningsvis er tekst fra HRCS-veiledningen limt inn.

1. Blood

Haematological diseases, anaemia, clotting and normal development and function of platelets and erythrocytes.

Advice:

Not all haemopoietic cells or 'blood cells' are erythrocytes. This term is often used to encompass lymphocytes, macrophages, granulocytes and other white cell types. Check the context of the research eg if the research relates to leukaemia it will be referring to lymphocytes etc and not to erythrocytes and therefore should not be coded as Blood.

Det er en Web of Science kategori som dekker feltet bra, nemlig HEMATOLOGY. Av definisjonen ovenfor ser en imidlertid at bare forskning omhandlende røde blodceller og blodplater er omfattet av kategorien. For å få et bedre samsvar med HRCS-kategorien bør derfor forskning på leukemi og hvite blodceller ekskluderes. Dette kan gjøres ved å benytte søkeord for dette.

Søkeord - inkluderes	Søkeord – ekskluderes
haematological disease, hematological disease, anaemia, anemi, platelet, erythrocyte, thrombocyte, hemolytic, haemolytic, favism, hemoglobinuria, haemoglobinuria, hemostasis, haemostasis, thrombocyte, serum albumin, haematopoiesis, hematopoiesis, haemopoiesis, hemopoiesis, erythropoiesis, thrombosis, thrombo, polycythemia, polycythaemia, polyglobulia, thrombocytosis, hemophili, haemophili, cryoprecipitate, Hemochromatosis, Haemochromatosis	lymphocyte, macrophage, granulocyte, leukaemia, leukemia, leukocyte, agranulocyte, monocyte

2. Cancer

All types of cancers (includes leukaemia)

Advice:

Do not code to the site of the cancer. However if the research involves studying a condition that predisposes to cancer then it may be appropriate to code for this condition as well. Eg The role of Barrett's oesophagus in cancer would be 50% Oral and Gastrointestinal and 50% Cancer.

Similarly research on pathogens associated with the development of cancer should be coded 50% Cancer 50% Infection Studies of the normal role of oncogenes, tumour suppressor genes, cell cycle checkpoints in a non diseased cell should be coded as 50% Generic and 50% Cancer.

Excludes general studies of angiogenesis which should be coded as Cardiovascular. However the development of anti-angiogenic drugs to inhibit tumour growth would be coded as Cancer.
Excludes normal studies of cell cycle and DNA replication and repair which should be coded as Generic.

Klassifiseringssystemet åpner for at artikler kan klassifiseres i flere enn én HRCS-kategori. Grunnprinsippet er imidlertid at en artikkel bare skal klassifiseres under den sykdomsgruppen den tilhører. Det innebærer at for eksempel studier av lungekreft skal klassifiseres under 2. Cancer og ikke under 17. Respiratory. Det skal således normalt ikke brukes dobbelklassifiseringer for kreft.

Noen av prinsippene beskrevet ovenfor er imidlertid vanskelig å justere for ved en automatisk identifisering, for eksempel at tilstander som predisponerer for kreft skal kodes dobbelt.

Søkeord - inkluderes	Søkeord – ekskluderes
cancer, oncogene, malignant neoplasm, metastasis, metastases, carcinogen, radiation therapy, radiotherapy, oncology, radioimmunotherapy, chemotherapy, sarcoma, carcinoma, chemoimmunotherapy, carcinogenesis, oncogenesis, malignant tumours, malignant tumors, malignant cell, leukemia, cancerous	

3. Cardiovascular

Coronary heart disease, diseases of the vasculature and circulation including the lymphatic system, and normal development and function of the cardiovascular system

Advice:

Includes general circulation research, vasculitis and general angiogenesis studies.

Includes atherosclerosis and pulmonary hypertension.

Includes congenital heart disorders.

Excludes studies investigating blood flow to the brain which should be coded as Stroke.

Excludes angiogenesis studies which relate to a diseased state eg development of anti-angiogenic drugs to inhibit tumour growth, which would be coded as Cancer.

Søkeordene listet nedenfor identifiserer mange artikler, også noen som kan synes mer perifere for feltet. Det kan derfor være en viss overestimering.

Søkeord - inkluderes	Søkeord – ekskluderes
coronary heart disease, vasculitis, pulmonary hypertension, congenital heart disorder, congenital heart disease, vascular, cardiovascular, angiology, circulatory system, arteries, artery, arterial, blood pressure, capillaries, aorta, aortic, coronary artery disease, angina, myocardial infarction, cardiomyopathy, artery, arterial, cardiomyopathy, cardiac, ischemia, ischemic, myocardium, arteriosclerosis, atherosclerosis, heart failure, heart defect, cardiomyocytes, myocardiocytes, ventricle, ventricular, hypercholesterolemia, dyslipidemia, veins, venous, atrial fibrillation, cardiopulmonary, myocardial, heart transplantation, cardiology, thrombosis, lymphatic, coronary interventions	Stroke, cerebrovascular accident, cerebrovascular insult, brain attack, cerebral infarction, brain ischemia, ischemic attack, cerebral hemorrhage, cubarachnoid hemorrhage, cerebral hypoperfusion, intercranial hemorrhages, cerebrovascular apoplexy, brain vascular accident

4. Congenital Disorders

Physical abnormalities and syndromes that are not associated with a single type of disease or condition including Down's syndrome and cystic fibrosis

Advice:

Includes physical malformation and congenital syndromes that are associated with multiple diseases and conditions.

Excludes inherited single disease disorders (even when referred to as 'congenital') which should be coded under the appropriate Health Category

Fagfeltet kan være vanskelig å identifisere basert på beskrivelsen ovenfor. Uansett er det svært få publikasjoner som vil omfattes av fagfeltet.

Søkeord - inkluderes	Søkeord – ekskluderes
congenital disorders, congenital diseases, congenital anomalies, Down's syndrome, Down syndrome, cystic fibrosis, teratogen, trisomy	

5. Ear

Deafness and normal ear development and function

Advice:

Includes normal studies on this organ or the condition of deafness.

Studies that involve the auditory nerve leading to the brain may be double coded as 50% Neurological and 50% Ear as appropriate.

Excludes mapping of hearing in regions of the brain in which case the Neurological category should be applied.

Søkeord - inkluderes	Søkeord – ekskluderes
deafness, ear, otologic, hearing, tinnitus, hyperacusis, otosclerosis, otoscleroses, tympanic, retrocochlear disease, otomycosis, otitis, otitides, myringoscleroses, tympanosclerosis, tympanoscleroses, earache, otalgia, otalgias, auditory performance, auditory system, auditory cortex	

6. Eye

Diseases of the eye and normal eye development and function

Advice:

Includes normal studies on this organ or the condition of blindness.

Studies that involve the optic nerve leading to the brain may be double coded as 50% Neurological and 50% Eye as appropriate.

Excludes mapping of vision in regions of the brain in which case the Neurological category should be applied.

Søkeord - inkluderes	Søkeord – ekskluderes
asthenopia, cogans syndrome, cornea, sclera, fibrous tunic, vascular tunic, uvea, conjunctiva, keratitis, blindness, retina, retinal, ectropion, entropion, dermatochalasis, macular degeneration, hyperopia, myopia, visual impairment, vision impairment, vision loss, contact lens, visual impairment, intraocular, oculi muscle, visual disability, visual disabilities, eye disease	

7. Infection

Diseases caused by pathogens, acquired immune deficiency syndrome, sexually transmitted infections and studies of infection and infectious agents.

Advice:

Infection should be used for all research on pathogens and diseases caused by infection. These studies should be coded as 100% Infection and should not also be coded to the target disease site eg infection by malaria parasite should be coded as 100% Infection and not Blood and respiratory infections should be coded as 100% Infection and not Respiratory

Studies of the downstream consequences of infection may also be coded to the disease site eg surgery on diseased liver from Hepatitis C infection can be coded as 50% Infection and 50% Oral and Gastrointestinal.

General studies on sexual health should be coded as 50% Infection and 50% Reproduction.

Studies of cancer causing pathogens should be coded 50% Cancer 50% Infection.

Sepsis should be coded 50% Infection and 50% Inflammatory and Immune.

Excludes natural tolerance and immunity to infections which is coded as Inflammatory and Immune System.

Excludes TSEs, CJD and prion studies which should be coded as Neurological regardless of whether the study is investigating transmission or mechanism of disease.

Dette fagfeltet er noe vanskelig å identifisere. Det er kun brukt noen få helt generelle søketermer for å avgrense fagfeltet. Men det er i tillegg en egen kategori for dette INFECTIOUS DISEASES.

Søkeord - inkluderer	Søkeord – ekskluderer
infection, infectious, pathogen, antibiotic, antiviral	transmissible spongiform encephalopathies, Bovine Spongiform Encephalopathy, Creutzfeldt-Jakob Disease, prion protein

8. Inflammatory and Immune System

Rheumatoid arthritis, connective tissue diseases, autoimmune diseases, allergies and normal development and function of the immune system

Advice:

Includes listed immune diseases and studies of normal immune response.

Also includes natural tolerance and immunity to infections.

Sepsis should be coded 50% Infection and 50% Inflammatory and Immune System.

Excludes asthma which should be coded as Respiratory. However studies of allergies in asthma sufferers can be coded as 50% Inflammatory and Immune System for allergies and 50% Respiratory for asthma.

Most diseases are associated with an inflammatory response. Code to the disease under investigation and not to the accompanying inflammatory response.

Excludes atherosclerosis which is classified as Cardiovascular regardless of whether the investigation involves immune cells.

Søkeord - inkluderer	Søkeord – ekskluderer
immune, immunity, immunological, rheumatoid arthritis, autoimmune, allergies, allergy, inflammation, myopathies, leukocytes, phagocytes, macrophages, neutrophils, dendritic cells, innate lymphoid cells, mast cells, eosinophils, basophils, natural killer cells, lymphocytes	atherosclerosis, asthma

9. Injuries and Accidents

Fractures, poisoning and burns

Advice:

This category is designed to capture research directly relating to external accidents, injuries and physical trauma and should not include endogenous studies of damage such as ischemic injury.

Intervention studies preventing or treating falls should be coded to Injuries and Accidents. It may be appropriate to also code for specific conditions if the study involves a particular patient group.

Excludes studies of wound healing, unless the wound was acquired by accident.

Søkeord - inkluderer	Søkeord – ekskluderer
major trauma, blunt trauma, non-penetrating trauma, blunt force trauma, abdominal trauma, physical trauma, penetrating trauma, perforating trauma, broken bone, bone fracture, gunshot, burn, poisoning, drug overdose, vehicle crashes, drowning, barotrauma, penetrating wound, nonpenetrating wound, wound Infections, frostbite, heat stress disorder, heat stress syndrome, heat cramp, cartilage fracture, accidental electrocution, traumatic amputation	

10. Mental Health

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

Advice:

Mental Health should be used for all normal cognitive research categorised to 1.2 Psychological i.e. behaviour learning, memory, language, perception etc.

Studies dealing with the brain of individuals with a psychological condition listed in the Mental Health category should be coded as 100% Mental Health and not as Neurological.

Fagfeltet er tilfredsstillende dekket av den tidsskriftsbaserte avgrensningen, og det inkluderer bare noen få supplerende søkeord.

Søkeord - inkluderer	Søkeord – ekskluderer
depression, schizophrenia, psychosis, personality disorders, suicide, anxiety, eating disorder, learning disabilities, autistic spectrum disorders	

11. Metabolic and Endocrine

Diabetes, thyroid disease, metabolic disorders and normal metabolism and endocrine development and function

Advice:

This category includes basic studies of metabolism but not generic signalling pathways involving kinases.

Studies on nutrition, diet and obesity are context based and should only be coded as Metabolic and Endocrine if they relate to metabolism.

Also includes studies of hormone receptors.

Excludes reproductive hormones usually coded as Reproduction.

Søkeord - inkluderer	Søkeord – ekskluderer
diabetes, diabetic, thyroid, metabolic, metabolism, endocrine, gland, catabolism, anabolism, hypothalamus, gastrointestinal tract, insulin	

12. Musculoskeletal

Osteoporosis, osteoarthritis, muscular and skeletal disorders and normal musculoskeletal and cartilage development and function

Advice:

Includes research into limbs and the face.

Søkeord - inkluderer	Søkeord – ekskluderer
osteoporosis, osteoarthritis, skeletal, musculoskeletal, cartilage, tendon, ligament, cartilage, skeleton, bursae, bursa arthroplasty, arthroplasties, arthrosis, osteoarthritis, rheumatic, arthritis, prosthesis, rheumatoid	bone marrow

13. Neurological

Dementias, transmissible spongiform encephalopathies, Parkinson's disease, neurodegenerative diseases, Alzheimer's disease, epilepsy, multiple sclerosis and studies of the normal brain and nervous system

Advice:

This category is suitable for studies of the 'wiring' of the brain.

Includes studies of TSEs, CJD and prion protein.

Studies on circadian rhythm are often but not always Neurological.

Studies on headaches and migraine are often but not always Neurological.

Excludes studies dealing with the brain of individuals with a psychological condition listed in the Mental Health category which should be coded as 100% Mental Health.

Søkeord - inkluderer	Søkeord – ekskluderer
Dementia, Parkinson's disease, neurodegenerative disease, Alzheimer's disease, epilepsy, multiple sclerosis, migraine, transmissible spongiform encephalopathies, Bovine Spongiform Encephalopathy, Creutzfeldt-Jakob Disease, prion protein, neurological, neurodevelopmental, neurogenetic, neuropsychological, neuroborreliosis, neurons, neuropathic, neuropathy, neuropharmacological, neurocognitive, hippocampus, hippocampal, cerebral, spinal cord, nervous system, cortical, neural, intracranial, cerebrospinal	

14. Oral and Gastrointestinal

Inflammatory bowel disease, Crohn's disease, diseases of the mouth, teeth, oesophagus, digestive system including liver and colon, and normal oral and gastrointestinal development and function

Advice:

Includes dental research and studies on the mouth, throat, stomach, liver, pancreas (not diabetes/insulin related), gut and colon. Also suitable for problems linked to food absorption by the gut.

Søkeord - inkluderer	Søkeord – ekskluderer
inflammatory bowel disease, Crohn's disease, mouth, teeth, oesophagus, digestive system, liver, colon, gastrointestinal, dental, throat, stomach, pancreas, gut, colon odontal, periodontal, caries, dentist, oral health, pancreatic endoscopic, ulcerative colitis	diabetes, diabetic, insulin

15. Renal and Urogenital

Kidney disease, pelvic inflammatory disease, renal and genital disorders, and normal development and function of male and female renal and urogenital system

Advice:

Includes female and male sexual organs and associated diseases where the research is not directly linked to reproductive issues such as developmental studies.

Søkeord - inkluderer	Søkeord – ekskluderer
kidney, pelvic inflammatory disease, renal, urogenital, renin, nephritic, nephrotic, urinary, dialysis, nephrons, genitourinary, ureters, bladder, urethra, urination, urine, urethra, ureter, urea, uric acid	

16. Reproductive Health and Childbirth

Fertility, contraception, abortion, *in vitro* fertilisation, pregnancy, mammary gland development, menstruation and menopause, breast feeding, antenatal care, childbirth and complications of newborns

Advice:

Includes studies of the effects of exposure to factors while in utero on the foetus or new borns.

Includes studies of ante and post natal care such as issues affecting new borns.

Studies of sexual health should be coded 50% Reproductive Health and Childbirth and 50% Infection.

May include studies of reproductive hormones.

If the study is about the long term effects on children or adults, where the foetus is not involved in the investigation, the Reproductive Health and Childbirth category should not be used.

It also excludes basic developmental studies in non-human models.

Søkeord - inkluderer	Søkeord – ekskluderer
fertility, contraception, abortion, <i>in vitro</i> fertilisation, pregnancy, mammary gland, menstruation, menopause, breast feeding, antenatal care, childbirth, newborn, new borns, reproductive health, sexual health, birth control, childbearing, genital, maternal health, gynecologic, gynecology, placenta, afterbirth, pregnant, maternal age, postmenopausal, obstetrics, birth weight, stillbirth	

17. Respiratory

Asthma, chronic obstructive pulmonary disease, respiratory diseases and normal development and function of the respiratory system

Advice:

Studies of allergies in asthma sufferers can be coded as 50% Inflammatory for allergies and 50% Respiratory for asthma.

Excludes pulmonary hypertension which should be coded as Cardiovascular.

Excludes respiratory infections which should be coded as Infection.

Søkeord - inkluderer	Søkeord – ekskluderer
asthma, chronic obstructive pulmonary disease, respiratory, respiration, lung, breathing, trachea, bronchi, bronchioles, diaphragm, alveoli	

18. Skin

Dermatological conditions and normal skin development and function

Advice:

Can include studies of ulcer wound healing.

Søkeord - inkluderer	Søkeord – ekskluderer
dermatological, melanin, melanocytes, mesodermal epidermis, dermis, hypodermis, keratin, psoriasis, dermatitis, skin hypoxia, skin disease, acne, dermatosis, actinic keratosis, actinic keratoses	

19. Stroke

Ischaemic and haemorrhagic

Advice:

Only studies investigating blood flow to the brain should be coded as Stroke.

General blood circulation and cardiovascular disease research should be coded as Cardiovascular.

Søkeord - inkluderer	Søkeord – ekskluderer
stroke, cerebrovascular accident, cerebrovascular insult, brain attack, cerebral infarction, brain ischemia, ischemic attack, cerebral hemorrhage, subarachnoid hemorrhage, cerebral hypoperfusion, intercranial hemorrhages, cerebrovascular apoplexy, brain vascular accident	

20. Generic Health Relevance

Research applicable to all diseases and conditions or to general health and well-being of individuals. Public health research, epidemiology and health services research that is not focused on specific conditions. Underpinning biological, psychosocial, economic or methodological studies that are not specific to individual diseases or conditions

Advice:

The Generic category refers to research that is relevant to all diseases and conditions or to general health and well-being.

It is applicable to any research that can not be attributed to a particular disease or condition or to normal function of a specific type of cell or system.

If research is judged relevant to more than five Health Categories this category should be used.

If the main focus of the research is directed at several specified diseases and also has implications for many other conditions, the appropriate specific Health Categories should be used as well as applying the Generic category. (This does not apply to diseases that may be listed within the background information or are noted as 'being relevant' to the study under investigation - see the coding guidance on assigning Health Categories.)

Denne kategorien, som blant annet omfatter samfunnsmedisinsk forskning og epidemiologi, dekkes av flere ulike WoS-kategorier. Det vurderes at fagfeltet er tilstrekkelig avgrenset basert på den tidsskriftsbaserte identifiseringen slik at det ikke er behov for supplerende søkeord.

21. Other

Conditions of unknown or disputed aetiology (such as chronic fatigue syndrome/myalgic encephalomyelitis), or research that is not of generic health relevance and not applicable to specific health categories listed above

Advice:

The Other category is used for specific conditions or diseases that cannot be attributed to any of the named categories. This includes conditions of unknown or disputed aetiology e.g. chronic fatigue syndrome/myalgic encephalomyelitis.

An example of a condition of unknown aetiology that may be suitable to include is Gulf war syndrome.

Studies into animal welfare may be coded to Other.

Other also includes some types of social services research for specific 'healthy' at risk groups that is not of relevance to the general population e.g. research into services for young people at risk of domestic violence.

Søkeord - inkluderer	Søkeord – ekskluderer
chronic fatigue syndrome, myalgic encephalomyelitis	

Det er uklart hvilke søkeord som vil være aktuelle for denne kategorien, ut over de to søkeordene som er listet. I systemet som er foreslått, er ikke veterinærmedisin inkludert slik at «animal welfare» faller ut. Det er imidlertid et betydelig antall artikler, primært innenfor basal biomedisin, som ikke vil fanges opp av klassifiseringssystemet. Dette dreier seg for eksempel om generelle cellebiologiske studier, som ikke har direkte klinisk relevans eller er knyttet til bestemte sykdommer. Som nevnt er disse ikke inkludert. Skulle de vært inkludert, hadde det vært mest rimelig under 20. *Generic Health Relevance* (jf. definisjonen: “Underpinning biological, psychosocial, economic or methodological studies that are not specific to individual diseases or conditions”).

3 Prosedyre

Basert på gjennomgangen i kapittel 2, beskrives her den foreslåtte metodologien for HRCS-klassifisering trinn for trinn.

- 1) Populasjonen av artikler som skal analyseres identifiseres. Avgrensningen gjøres ved hjelp av WoS-kategoriene som dekker medisin og helsefag. Alle artikler som er klassifisert innenfor disse kategoriene inkluderes. I tillegg inkluderes artikler som er publisert i multidisiplinære tidsskrifter (for eksempel *Plos One* og *Science*). Sistnevnte artikler er fagklassifisert per artikkel i WoS, derfor inkluderes aktuelle artikler innenfor medisin og helsefag fra disse tidsskriftene. Analysen bør normalt begrenses til publikasjoner av typene *articles*, *reviews* og *letters*. Dette utgjør så *grunnsettet* av artikler som skal HRCS-klassifiseres. I en videreutvikling av systemet kan det suppleres med data fra andre datakilder enn WoS, slik at øvrig publisering også dekkes.
- 2) Excel-filen med HRCS-klassifisering av tidsskrifter benyttes for å tildele HRCS-kategorier til enkeltartikler. Slik vil hver artikkel i tidsskrifter som er klassifisert etter HRCS-systemet få en spesifikk klassifisering.
- 3) Det gjennomføres en søkeordbasert klassifisering som omfatter hele grunnsettet av artikler, også dem som allerede er klassifisert under trinn 2. Det søkes både i titlene og sammendraget til artiklene.
- 4) Det lages en samlefil hvor alle artikler som er indentifisert i trinn 2 og 3 inkluderes. Artikler som inneholder ekskluderende termer (for de kategoriene hvor dette er aktuelt) utelates fra de aktuelle HRCS-kategoriene. Også her søkes både i titlene og sammendraget til artiklene
- 5) Mange artikler vil være klassifisert i flere enn én HRCS-kategori. Som beskrevet i kapittel 2, skal det ikke brukes dobbelklassifiseringer for kreft. For artiklene som er indentifisert innen kreft, basert på søkeord eller tidsskrift, ekskluderes derfor dobbelklassifiseringer, med unntak av dobbelklassifiseringer i HRCS-kategorier som ikke er knyttet til spesifikke kroppsdelene: 4. Congenital Disorders 7. Infection 8. Inflammatory and Immune System 9. Injuries and Accidents, 10. Mental Health 20. Generic Health Relevance.
- 6) Ikke alle artikler i grunnsettet vil være klassifisert gjennom trinn 2 og 3. Det er en restkategori som må analyseres videre. For gjenværende artikler som er klassifisert innenfor WoS-kategoriene Sport Sciences og Nutrition & dietetics benyttes en standardisert HRCS-

klassifisering i henhold til fordelingen beskrevet i HRCS-veiledningen: 2. Cancer, 3. Cardiovascular, 11. Metabolic and Endocrine, 14. Oral and Gastrointestinal og 19. Stroke.³

- 7) Øvrige artikler i restkategorien inkluderes ikke, siden disse ikke lar seg klassifisere på relevante HRCS-kategorier med metoden. I en videreutvikling av systemet, kan det imidlertid vurderes å utvikle metodologi for å inkludere flere av disse, men mange av artiklene synes å falle utenfor medisin og helsefag.
- 8) For hver artikkel lages det en opptelling av antall HRCS-kategorier den er klassifisert i. Deretter fraksjoniseres artikkeltallene. En artikkel som for eksempel er klassifisert i fire forskjellige HRCS-kategorier gis en vektning på 0,25 artikkel i hver av kategoriene.
- 9) På bakgrunn av klassifiseringen kan ulike typer bibliometriske indikatorer utarbeides. Videre kan det beregnes indikatorer over forholdet mellom ressursinnsats og resultater i form av vitenskapelig publisering. I sistnevnte beregninger kan det være relevant å justere for innslaget av utenlandsk medforfatterskap (fraksjonalisering) siden et flertall av artiklene også har medforfattere fra andre land.

³ Jf: "Use the following standardised breakdown of health categories and percentages for general studies on nutrition, obesity or poor diet where there is no other indication of disease area: Cancer: 20% Cardiovascular 20%, Metabolic and Endocrine 20%, Oral and Gastrointestinal 20%, Stroke 20%."

4 Testresultater

Systemet beskrevet ovenfor ble testet ut på samlingen av norske artikler innen medisin og helse for perioden 2011-2015 og som er indeksert i Web of Science.⁴ Til sammen 24758 artikler ble identifisert gjennom systemet. Figur 1 viser hvordan artiklene fordelte seg på HRCS-kategorier. To indikatorer er beregnet, fraksjonalisert og ikke-fraksjonalisert fordeling. HRCS systemet legger opp til fraksjonaliserte fordelinger, slik at et prosjekt (eller en artikkel) som er henført til for eksempel tre ulike HRCS-kategorier, skal telle som 1/3 i hver av kategoriene. Det er imidlertid interessant i tillegg å beregne ufraksjonalisert andeler, men summen av de ulike kategoriernes andeler vil her langt overstige 100 prosent.

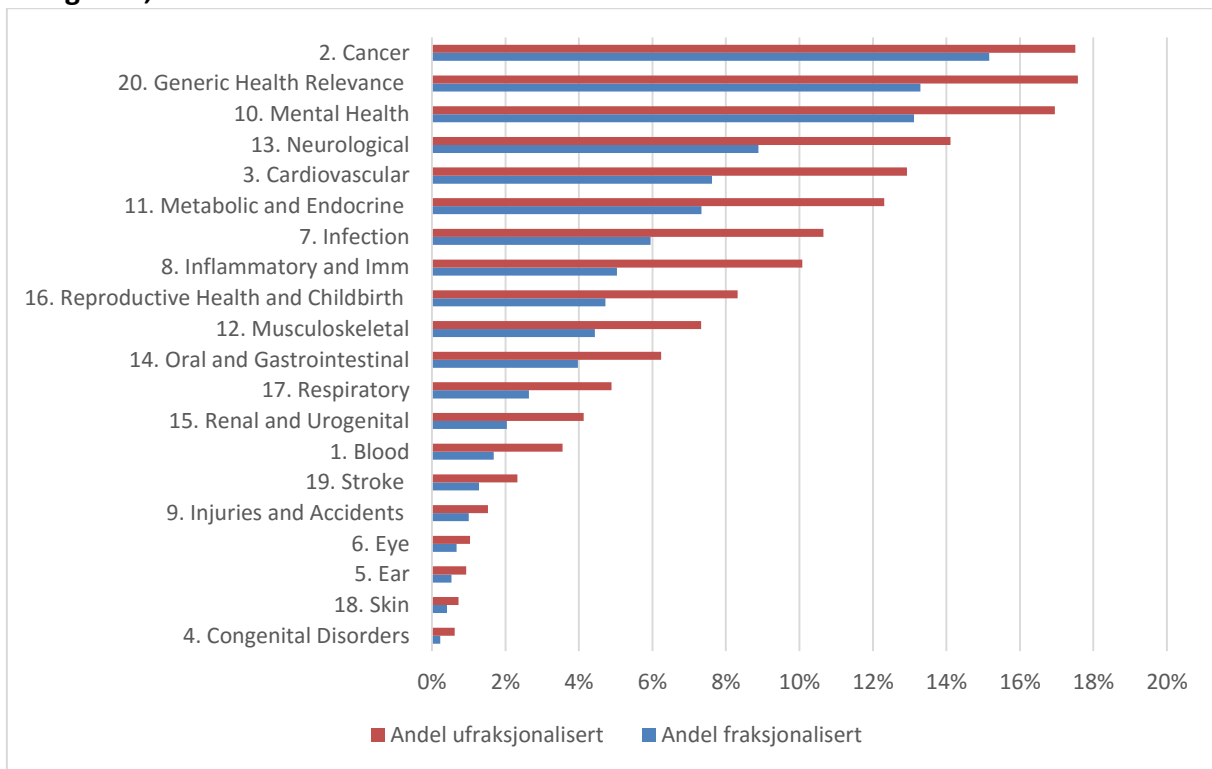
Som forventet, er fordelingen av publiseringen på HRCS-kategorier svært skjevfordelt hvor noen kategorier står for et betydelig antall artikler, mens andre er små i artikkelvolum. Kreftforskning/Cancer står for den største andelen med 15 prosent eller 18 prosent målt ufraksjonalisert. Det siste tallet kan tolkes som at 18 prosent av den vitenskapelige publiseringen innen norsk medisinsk og helsefaglig publisering var relatert til kreft.

Figur 2 er et bilde som er hentet fra det siste forskningsbarometeret (2016).⁵ Den viser omfanget av prosjektfinansiert helseforskning i Norge i 2015 fordelt på HRCS-kategorier (prosjekter finansiert av Forskningsrådet, de regionale helseforetakene, samt prosjekter med norsk deltakelse finansiert av EUs forskningsprogram Horisont 2020). En betydelig del av forskningen finansieres av helseforetakenes og forskningsinstitusjonenes grunnbevilgninger, og dette er ikke inkludert i figuren. Fagprofilen er dermed ikke nødvendigvis representativ for totalinnsatsen. Likevel er det interessant å se at mønsteret her i stor grad samsvarer med resultatene fra publiseringsanalysen. Dette tyder på at metodologien utviklet i prosjektet gir rimelige resultater. En forskjell det er verdt å merke seg er likevel at forskning på kreft og forskning med generell helserelevans synes å utgjøre større andeler av finansieringen enn av den vitenskapelige publiseringen.

⁴ Etter at denne testen ble gjennomført, ble det foretatt noen justeringer av systemet. Endringene er imidlertid såpass små at den overordnede profilen som presenteres her i liten grad vil bli påvirket.

⁵ Se <https://www.regjeringen.no/contentassets/d931558128774ac8932cf41033eb1c9e/forskningsbarometeret-2016.pdf>

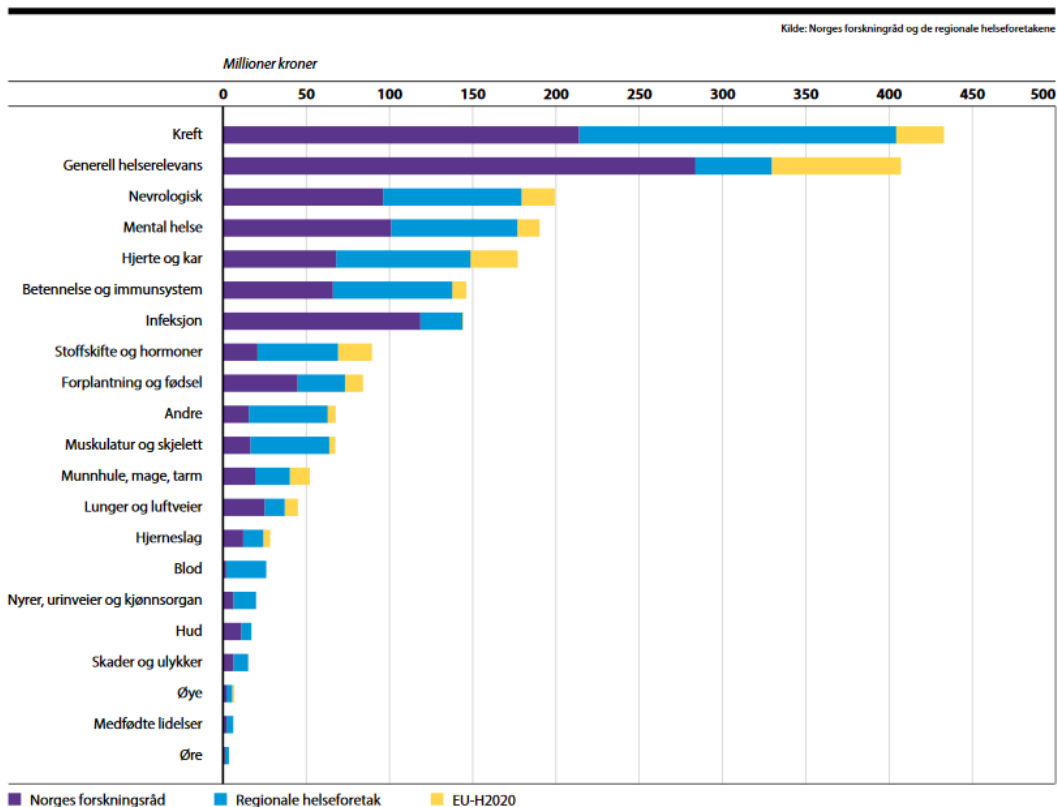
Figur 1. Test av fordelingen av norske artikler innen medisin og helse etter HRCS-kategorier, 2011-2015.



Figur 2. Prosjektfinansiert helseforskning I Norge etter HRCS-kategorier og finansieringskilde, 2015.

34

Prosjektfinansiert helseforskning i Norge etter helsekategorier (HRCS) og finansieringskilde 2015, millioner kroner



Figur 3 er et tilsvarende bilde som er hentet fra den siste analyserapporten til HRCS i Storbritannia.⁶ Den viser fordelingen av bevilgningene til landets medisinske og helsevitenskapelige forskning. Det er relativt stort samsvar i rangeringsrekkefølgen sammenlignet med den norske. Imidlertid synes fordelingen å være enda mer skjevfordelt enn den norske, og det er en del avvik for enkelte av kategoriene, blant annet for 10. Mental Health. Den britiske fordelingen kan imidlertid ikke brukes som validering av systemet foreslått her, siden Norge har en annen fagprofil enn Storbritannia.

Figur 3. UK Health Research Analyses, covering the 2014 calendar year, expanded the analysis to include 64 UK charity and public health research funders

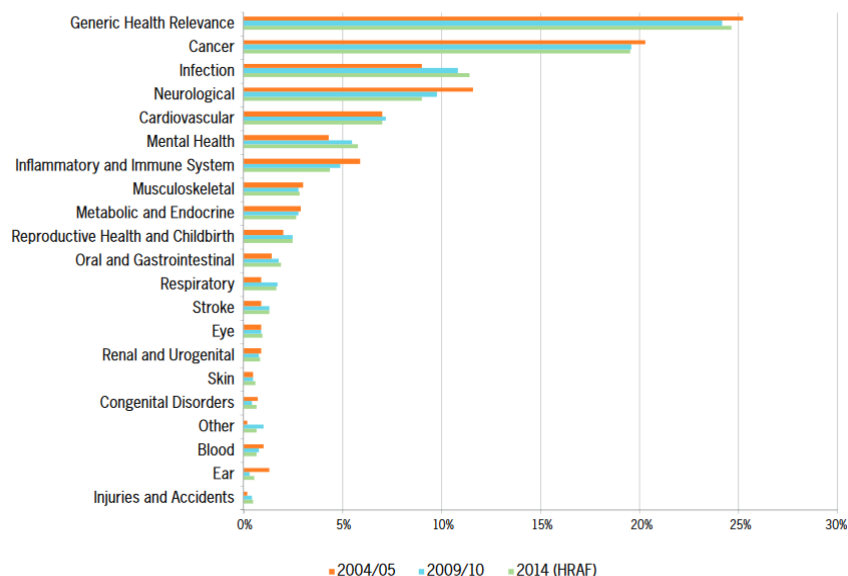


Figure 9 – Proportion of combined spend by health category for 2004/05, 2009/10 and 2014 (HRAF, 12 funders).

⁶ Se <http://www.hrcsonline.net/pages/uk-health-research-analysis-2014>

Vedlegg – oversikt over HRCS-klassifisering av tidsskrifter.

Oversikten foreligger som separat Excel-fil.

Nordisk institutt for studier av
innovasjon, forskning og utdanning

Nordic Institute for Studies in
Innovation, Research and Education

www.nifu.no