A complete overview of the PhD theses within the field of medicine and health science in Norway in 2018 by using the Health Research Classification System (HRCS)

Helen E Kvalem (✉ hekv@forskningsradet.no)

The Research Council of Norway   https://orcid.org/0000-0002-8289-9894

Frode Hovland Søreide

The Research Council of Norway

Bo Sarpebakken

Nordic Institute for Studies in Innovation

Dyveke L Hetland

The Research Council of Norway

Research

A complete overview of the PhD theses within the field of medicine and health science in Norway in 2018 by using the Health Research Classification System (HRCS)

Kvalem HE[1], Søreide FH[2], Sarpebakken B[3], Hetland DL[4]

[1] The Research Council of Norway, Dept. for Health Research and Health Innovation, [2]The Research Council of Norway, Dept. for Statistics and Evaluation, [3]Nordic Institute for Studies in Innovation, Research and Education, [4]The Research Council of Norway, Dept. for Open Researcher Arena

Summary

Background: The two-dimensional Health Research Classification System (HRCS) describes health research by the type of research undertaken (Research Activity, RA) and by the health issue or disease addressed (Health Category, HC). This is the first time HRCS has been used to classify PhD theses. **Material and methods**: All 485 PhD theses within medicine and health in Norway from 2018 were coded with HRCS. **Results**: *Cancer and Neoplasms* (12.1%), *Cardiovascular* (10.7%), *Mental Health* (10.5%) and *Generic Health Relevance* (9.8%) were the largest Health Categories whereas *Aetiology* (32.1%), *Evaluation of Treatment* (19.7%) and *Detection and Diagnosis* (13.3%) were the largest Research Activity categories. **Interpretation**: There is not a perfect overlap in HRCS profiles between the new PhDs in 2018 and the projects awarded from the main research funding organisations in Norway[1]. In terms of Research Activities, the disparity between HRCS coded PhDs and HRCS-coded research projects is greatest for RAs: A*etiology* (higher for PhDs) and *Prevention* (lower). While some major health challenges in Norway in DALYs like *Cancer and Neoplasms, Mental Health* and *Cardiovascular* are accordingly addressed in the PhDs 2018, the Health Categories *Musculoskeletal*, *Respiratory* and *Injuries and Accidents* are not.

Introduction

The Health Research Classification System was originally designed by the UK Clinical Research Collaboration (UKCRC) to provide strategic overview of health research funding, through the classification of awarded research projects (https://hrcsonline.net/). In Norway the Research Council (RCN), The Regional Health Authorities and some other health research funding organisations use the system to classify the projects they fund. The RCN has also initiated two pilot projects in order to gain experience with HRCS classification of scientific publications and full-time equivalents (FTEs) of researchers. (NIFU 2016:12, NIFU 2019:10). HRCS is a system used to classify research into eight main categories of Research Activity (e.g. *Underpinning*, *Aetiology*) and 21 Health Categories (e.g. *Blood*, *Cancer and Neoplasms*, *Cardiovascular)*, for overview of RA and HC see Supplementary Table 1 and 2. HRCS can be used on all research with focus on human health irrespective of academic disciplines. HRCS classification gives valuable insight when monitoring health research, especially for policy makers and research leaders. A PhD is the highest achievable academic degree and the topic of the thesis is an indicator of new research resources and prioritization within health and medical research. The HO21-monitor is a platform to provide an overview of Norwegian health research

---

(https://www.helseomsorg21monitor.no/). The present project gives valuable insight into what kind of health research – in terms of research type and disease focus – that is performed in Norwegian PhD theses. The main purpose of this study is to raise awareness and give an overview of the PhDs in 2018 within the field of medicine and health in Norway. A second purpose is to strengthen the value of the HO21-monitor as an information platform.

Material and methods

A list of the titles, names and affiliation for all 487 PhD theses from 2018 was received as an excel-file from the Nordic Institute for Studies in Innovation, Research and Education (NIFU) April 2019. All PhD theses summaries, full text thesis or included papers were collected and then coded manually during April-October 2019 by one researcher using HRCS. Out of the 487 PhD thesis 2 fell outside of the classification system with no outcome relevant for human health. We allowed for a maximum of two research activities and five health categories as recommended by the HRCS guidelines developed for research projects. All coding was performed using these guidelines as a standard. The PhD theses were coded and weighted, so that if the dimension Research Activity were split 50:50 and the dimension Health Category were split 20:20:20:20:20, the Health Category would contribute 10:10:10:10:10 towards each of the (50%) Research Activities. Universities with fewer than 15 PhD candidates were collapsed into one group to increase statistical power and to protect privacy and identity. The data handling and statistical analyses were performed with Excel and SPSS 26, using non-parametric tests, testing two-sided with significance level 0.05.

Results

HRCS distribution

The 485 PhD theses were coded, and all 8 main Research Activity codes and 21 Health Categories were in use. When weighting all combinations in a heatmap, the most frequent combinations of Research Activity and Health Category are marked in a deep green colour (Table 1).

The Health Category percentage ranged from 12.1% for *Cancer and Neoplasms* to 0.2% for *Congenital Disorders*. In order of magnitude *Cancer and Neoplasms* was followed by *Cardiovascular* (10.7%), *Mental Health* (10.5%) and *Generic Health Relevance* (9.8%) as the largest Health Categories. For Research Activities *Aetiology* (32.1%), *Treatment Evaluation* (19.7%) and *Detection and Diagnosis* (13.3%) were the largest categories followed by *Disease Management* (9.5%), *Treatment Development* (9.3%), *Underpinning* (7.1%), *Health Services* (6.4%) and finally *Prevention* (of diseases and conditions, and promotion of well-being) (2.6%). The most frequent combinations of HC and RA were *Aetiology – Cardiovascular* (4.1%), *Aetiology – Reproductive Health and Childbirth* (3.5%) and *Disease Management – Mental Health* (3.5%).

Universities

The University of Oslo (UiO) contributed most with a 44% share of the PhD theses, followed by the Norwegian University of Science and Technology 20% (NTNU), the University of Bergen 18% (UiB) and the Arctic University of Norway 11% (UiT). The seven universities/university colleges (out of eleven) with the fewest PhDs contributed with only 5% to the total. UiO had the highest numbers of PhDs within *Cancer and Neoplasms*, *Cardiovascular*, *Generic Health Relevance* and *Mental Health*. Looking at the individual universities the highest relative contribution to the health categories were for UiO *Cardiovascular*, NTNU *Cancer and Neoplasms*, UiB *Metabolic and Endocrine*, UiT *Generic Health Relevance* and the aggregated group Other with *Mental Health* (Figure 1a).

The most frequent Research Activity in the PhDs was *Aetiology*, for all the four largest universities in Norway (UiO, NTNU, UiB and UiT). However, UiO had a slightly different profile, with a higher share of *Treatment Evaluation* than the other institutions. The aggregated group Other has the highest share of PhDs within *Disease Management*, but is difficult to interpret, as Other is a conglomerate of seven institutions (Figure 1b).

Gender and age

Out of the 485 theses from 11 universities in Norway, 300 (62%) were written by women, 185 by men. The mean age at dissertation was 40.4 years for all, 40.3 years for men and 40.5 years for women. The men and women had a slightly different age-distribution curve; the men a quite classic bell shape with a tale towards higher age, but for women the curve appeared a bit more broken, but with the same mean age (Figure 2).

Testing gender differences on Health Categories showed that PhDs within the categories *Musculoskeletal* and *Infection* were statistically significant higher in men than women (p= 0.038, 10.4% of all 185 men & vs 4.8% of all 300 women) and (p=0.026, 7.2% vs 3.4% women). While for *Disputed Aetiology*, more women than men worked within this category (women 1.8% and men 0.54%, p= 0.046). As for the Research Activities; *Treatment Evaluation* was the only one to differ between men and women (25.1% vs 16.3% p=0.008, higher in men). Figures in relative values within the Research Activity and Health Category (Figure 3 and 4) and absolute values (not shown) will be easily accessible to explore further at https://www.helseomsorg21monitor.no/.

There seemed to be a change in pattern of Research Activity across age categories. We divided the PhD candidates in four age groups: "under 35", "35.1-45", "45.1-55" and "55.1+". For most RAs the distribution had a peak of PhDs in the second age category (35.1-45 years). After that we observed a general decrease in number of PhDs with increasing age category. However, for the use of the RAs *Underpinning* and *Treatment Development* this peak was in the youngest age category (p<0.001 and =0.002). On the other hand, the RAs *Treatment Evaluation* and *Disease Management* had a distribution skewed towards a higher age at PhD defence than the other RAs, (both p<0.001) (Figure 5).

We further investigated the Health Categories across the four age categories: the only proportion of PhDs changing with age groups appeared to be within *Mental Health* (1.3%, 4.9%, 1.5%, 2.7% of total HCs, p<0.001) (data not shown).

Nationality

Out of the 485 PhDs, 348 had Norwegian citizenship while 74 came from Europe, 38 from Asia, 19 from Africa and 6 from America and Oceania. Out of the 348 Norwegians, 49 (14%) received their basic education qualifying for a PhD in another country than Norway (Table 2). In terms of Health Categories, we only observed a difference for PhDs falling into the Health Category *Skin*; with more non-Norwegian citizenship candidates (p=0.006). The PhD theses from Norwegian citizens more often fell into the Research Activities *Treatment Evaluation* (Norwegians 23% vs non-Norwegians 11% , p=0.010) and *Disease Management* (Norwegians 11% vs non-Norwegians 6%, p=0.047) but less often in *Underpinning* (Norwegians 5% vs non-Norwegians 12%, p=0.020).

Discussion

The main finding in this study is that the categories *Aetiology* (32.1%), *Evaluation of Treatment* (19.7%) and *Detection and Diagnosis* (13.3%) were the largest Research Activities and *Prevention* (2.6%) the smallest, when classifying PhDs. While for the Health Categories; *Cancer and Neoplasms*

(12.1%), *Cardiovascular* (10.7%), *Mental Health* (10.5%) and *Generic Health Relevance* (9.8%) were used the most.

Methodological issues

As the Health Research Classification System is designed to classify research projects, it is not given that it is directly applicable to PhDs. However, PhD-projects are indeed research projects, but with the limitations that they (usually or ideally) should be completed within a timeframe of 3-4 years, where the candidate is given qualified mentoring and completes an essential research curriculum. This time limitation may affect the choice of Research Activity. As the high number of PhDs within the category *Aetiology* may indicate, some research activities may either fit better with the PhD format, or that certain scientific communities are better at gaining funding for PhDs. This could be better routines for including PhDs in grant applications.

All manual coding involves a risk for human failure, as could be the case in this project. The challenges of coding were also highlighted as a source of error from a project coding research projects at Oslo Metropolitan University, using HRCS for strategic research planning (Jamtvedt 2/2018). We have tried to minimize coding inconsistency by letting the same researcher do all the coding, following the HRCS guidelines and by using a random sample for validation and adjustment after completion of the 485 PhDs. We do however recommend interpreting the results with caution, not because of the coding, but because of the relatively small sample size and the time frame of one year only. We strongly recommend repeated measurements the coming years to identify time trends and a general pattern for HRCS in PhDs.

Subgroups

Our data gave us the opportunity to explore the subgroups variables gender, university, citizenship and country for the general education qualifying for a PhD.

Universities

Across all universities *Aetiology* was the most dominating Research Activity. One explanation may be that PhD projects connected to the large health cohorts in Norway, like HUNT and The Norwegian Mother, Father and Child Cohort Study (MoBa), will fall into the *Aetiology* category. Another explanation may be that projects with *Aetiology* as Research Activity are more suitable for the timeframe of a PhD, but these are just speculations. However, the University of Oslo had a large proportion of PhDs within *Treatment Evaluation*. This is probably due to the expectedly large number of medical doctors working on their PhDs at the university hospitals affiliated with UiO.

Gender

The number of new doctorates within medicine and health sciences has increased significantly over the past 25 years, from about 100 in the early 1990s to close to 500 in 2018. Over the same period more and more women have completed their PhDs in the discipline. At the beginning of the 1990s about one quarter of the new doctorates were women, but the proportion of women increased sharply. In 2002 there was equal gender balance for the first time, and in the last seven years the share of women defending a PhD thesis within medicine and health has been around and above 60% ([NIFU, Doctoral degree register](#)). The uneven distribution of gender where more women than men complete a PhD within medicine and health is just as expected and in line with the time trends (Figure 6). The overall representation of males is approximately 2/5 and females 3/5 within all health and research categories in our study. The two exceptions were in Health Category *Infection* and the Research Activity *Treatment Evaluation*, with higher proportion of men than expected.

Nationality

Another trend is the increasing number of foreigners among doctoral candidates. In 2018, 28% had a foreign citizenship among those who obtained a PhD in medicine and health in Norway. This is the highest share so far. By comparison, foreigners accounted for less than 10% of the doctoral candidates in this field of subjects during the 1990s. Nonetheless, the proportion of foreign nationals among doctoral candidates is lower in medicine and health than in science and technology (NIFU, Doctoral degree register). In our study, we see a flux of PhD students coming to Norway from abroad. Out of the 485 PhD candidates 147 came with a basic education (master, medical doctor etc) from another country, probably moving to Norway for the PhD. However, these data cannot tell us about the flux the other way, i.e. how many Norwegians completes their PhD outside of Norway.

HRCS distribution - In line with other findings?

We investigated if the HRCS profiles of the PhDs were reflected in the HRCS profiles in a large number of research projects in Norway. These are the projects presented in the HO21-monitor funded by *The Research Council of Norway, The Norway Regional Health Authorities, The Norwegian Cancer Society* and *The European Union* across the years 2014-18 (www.helseomsorg21monitor.no). The order of the Research Activities is not random but indicates an evolvement from basic research to applied health services research. Categories 1 and 2 can be referred to as *Basic research*, 3 as *Prevention of diseases*, categories 4, 5 and 6 as *Translational and clinical research* and 7 and 8 as *Health service research*. When we compare these four summarized groups it is only *Prevention of diseases* that stands out as underrepresented in the PhDs 2018 compared with the HO21-monitor projects (2014-18) (Table 3). The comparison of the top three RA from Norwegian PhDs 2018, UK 2018 research grants and HO21-monitor projects 2014-18 listed in table 4, shows *Aetiology* is on top in all three (UK Clinical Research Collaboration, 2020) and in accordance with earlier findings from the UK (Carter et al. 2016).

When we compare the HRCS profile of the PhDs with the HO21-monitor projects, there is some overlap, but not entirely. We did not find any significant correlation between the Research Activities from HO21-monitor projects and PhDs 2018 (Spearman's rho correlation 0.35, p=0.396). The low percentage of PhDs within the RA *Prevention* (Figure 7) may be explained by the research grants within the field of vaccines, that generate PhDs abroad and are hence not registered in Norway. When comparing the HC distribution based on the PhDs and research funding, we found a strong correlation for the Health Categories; the correlation coefficient 0.91 (p<0.001) (Figure 8).

Burden of Disease

Figure 9 shows the share of PhDs and the share of the disease burden according to HRCS health categories in a correlation scatterplot. Although it is possible to use the HRCS system to compare with disease burden, the categories of the two systems do not directly overlap. It is not possible to link the Health Categories: *Inflammatory and Immune System*; *General Health Relevance*; and *Disputed Aetiology and Other* to a single disease. These categories are therefore excluded from this analysis. This current mapping of disease burden categories to HRCS health categories was conducted by the Research Council of Norway for the purpose of analyzing its own research portfolio (unpublished). Similar mappings have been performed previously (Kinge et al. 2014, Madsen and Aagard. 4/2019). Global burden of disease (GBD) is the name of a global disease burden study conducted by the Institute for Health Metrics and Evaluation at the University of Washington. The GBD study incorporates both the prevalence of a given illness or injury, and their relative health

effects - such as death and disability. The disease burden is measured in Disease Adjusted Life Years, also known as DALYs. The use of HRCS to compare the relation between DALYS and the resources allocated to medical research is elaborated in Madsen et al 2019. This Danish study shows that there is a discrepancy between the two i.e. due to the tendency that private and governmental awards are often given to the same HCs, while others are poorly funded.

The major health challenges in disability-adjusted life years (DALYs) in Norway (recoded into HRCS) are *Cancer and Neoplasms* (17%), *Mental Health* (13%), *Musculoskeletal* (12%) and *Cardiovascular* (10%) *Neurological (*9%) *Injuries and Accidents* (8%) (IHME, global burden of disease 2017). When comparing burden of disease with the distribution of PhDs across health categories, we find that the categories *Cancer and Neoplasms* (12,1%), *Cardiovascular* (10,7%) and *Mental Health* (10,5%) are reflecting the DALYs nicely. However, this does not seem to be the case for the topics *Musculoskeletal (7% PhD vs 12% DALY)*, *Respiratory (3% vs 5.2% DALY)* and *Injuries and Accidents (1% vs 8% DALY)* where we see the largest gaps between PhDs and DALYs.

In conclusion, this is the first study of its kind, and future studies will be able to show to what degree our findings are a result of yearly fluctuations in the topics of PhD theses.

References

https://hrcsonline.net/

Aksnes. Klassifisering av vitenskapelig publisering etter Health Research Classification System (HRCS). NIFU Arbeidsnotat 2016:12. https://nifu.brage.unit.no/nifu-xmlui/bitstream/handle/11250/2411668/NIFUarbeidsnotat2016-12.pdf?sequence=1&isAllowed=y

Wiig et al. HRCS-kategorisering av FoU-aktivitet I Norge I 2017. NIFU Arbeidsnotat 2019:10. https://nifu.brage.unit.no/nifu-xmlui/bitstream/handle/11250/2616413/NIFUarbeidsnotat2019-10.pdf?sequence=1&isAllowed=y

https://www.helseomsorg21monitor.no/

NIFU, Doctoral degree register (14.02.2020). http://www.foustatistikkbanken.no/nifu/

Jamtvedt. Erfaringer med å identifisere helseforskning. Forskningspolitikk 2/2018. http://fpol.no/wp-content/uploads/2018/06/Forskningspolitikk-2-2018.pdf

UK Health Research Analysis 2018 (UK Clinical Research Collaboration, 2020) ISBN 978-0-903730-29-7 https://hrcsonline.net/reports/analysis-reports/uk-health-research-analysis-2018/

Carter et al. Health Research Policy and Systems (2016) 14:48. DOI 10.1186/s12961-016-0116-1

Kinge et al. Health Research Policy and Systems 2014, 12:64 http://www.health-policy-systems.com/content/12/1/64

Madsen and Aagaard. Svag sammenhæng mellem samfundsbehov og finansiering af sygdomsrelateret forskning, Forskningspolitikk 4/2019. https://www.fpol.no/wp-content/uploads/2019/12/Forskningspolitikk_4_2019.pdf

IHME, global burden of disease 2017. https://vizhub.healthdata.org/gbd-compare/

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Availability of data and materials
The datasets during and/or analysed during the current study available from the corresponding author on reasonable request.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
HEK coded the PhDs using HRCS, performed the statistical analyses and wrote the manuscript. FHS gave valuable input throughout the process. BS provided the data, performed analyses on gender and citizenship and gave valuable input throughout the process. DLH is an experienced HRCS-coder and guided HEK through the coding process and gave valuable input throughout the process. All authors read and approved the final manuscript

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- HRCSandPhDsfigures.pdf