

厚生労働科学研究費補助金(地球保健政策課題推進研究事業)

「我が国の世界保健総会等における効果的なプレゼンスの確立に関する研究」(H29-地球保健一般002)

平成30年度分研究報告書

Context and challenges of Japan's health system

研究担者 坂元清香

東京大学大学院医学系研究科 国際保健政策学教室 特任研究員

研究協力者

渋谷健司

東京大学大学院医学系研究科 国際保健政策学教室 教授

研究要旨

UHC(すべての人に基本的な保健サービスを支払い可能な価格で普及させること)が大きな政策目標となったグローバルヘルス分野において、我が国の知見がアジア諸国を中心とした発展途上国から求められている。また、低成長と少子高齢化の中で多くの課題が噴出し、我が国がどのように対応していくかが世界の注目を集めている。UHCはWHO総会をはじめとして各種国際会議にて必出の議題となっており、また2019年にはUHCに関する国連ハイレベル会合の開催もされ、UHCに関する議論は今後も盛り上がる事が予想される。本研究は、WHO総会等の主要会合における日本のプレゼンス向上を大目標に掲げるものであるが、とりわけ、G7伊勢志摩サミット以降日本が牽引し、また今後国際的にも議論が盛り上がるであろうUHCに焦点を当て、UHCを推進する上で我が国の比較優位性を抽出するものである。主な研究目的は1) WHO Asia-Pacific Health Observatory(APO)の枠組みを活用し、我が国の保健医療制度の現状と課題及び将来像を、実証的かつ包括的に分析すること、2) Global Burden of Disease(GBD)の枠組みを用い、人口動態や疾病構造の劇的な変化が都道府県レベルでどのような影響を及ぼしているかを明らかにすることで、UHC達成に必要な不可欠な格差解消への示唆を得ることである。今年度は2) Global Burden of Disease(GBD)の枠組みを用いて保健医療制度の現状と課題を正確に理解する上で重要とされている主要死因データの的確性について、メルボン大学の協力の元で検証を行なった。国の疾病構造や主要死因を分析する上で、死因として不適切または使用不能な主要死因は、比較研究を行なった6カ国の死因データのうち18%であり、日本では25%であった。情報が不十分な死因データは6カ国全体の8%を占めており、日本では11%であった。例えば、日本の70歳以上の死因のうち、17%は「高齢」によるものと診断されていた。日本の死因データの約1/4が不適切であることは、我が国の疾病負荷を正しく把握し、適切な保健医療制度及び投資をする上で大きな課題となりうる。これらの結果はInternational Journal of Public Healthで発表した。

A . 研究目的

The Global Burden of Disease (GBD) studies have documented remarkable improvements in health that have occurred during the last decade, but also how unevenly health outcomes are distributed between and within populations. A recent study has raised concerns that adult death rates for many diseases have plateaued and, in some cases, increased, including in high-income countries (Roth et al. 2018). Subnational data have demonstrated surprising health inequality in some countries with well-developed health systems (AIHW 2018; Chammartin et al. 2016; Mahapatra et al. 2007; Roth and Dwyer-Lindgren 2017). To be able to monitor such health trends and the impact of interventions accurately, valid, reliable, regular and up-to date national mortality and morbidity data are essential (Lopez 2013; Shibuya 2006).

The global health goals and accountability for their achievements have led to significant interest in monitoring data quality and to the development of summary indicators such as the Vital Statistics Performance Index (VSPI) which measures the quality and timeliness of available mortality data (Mikkelsen et al. 2015; Philips et al. 2014). However, to be able to determine what actions need to be taken to improve statistical outputs, a more comprehensive review of the data is needed to better understand the main data quality issues and their origins. The usual research methodology to assess the accuracy of causes of death (COD)

is to undertake an independent review of a sample of medical records and compare the records with the cause(s) written on the death certificate (Alpe'rovitch et al. 2009), or to compare the clinical COD to an autopsy-based COD (Schdev 2001). Such studies, however, are complex and expensive to carry out and usually conducted in only one or a handful of hospitals. The findings have often indicated that even in countries with well-functioning civil registration and vital statistics (CRVS) systems, the quality of the medical certification is not as good as might be expected (Rampatige et al. 2014; Adair et al. 2019). Accuracy in COD certification is likely to be a growing issue in countries experiencing significant population ageing, in particular related to dementia and multiple chronic conditions that make accurate and consistent certification more challenging (Naghavi et al. 2010).

Within this context, it is crucial to be able to understand how well national mortality data systems of high-income countries are performing, given the expectations for them. To address this issue, we assessed the certification specificity and policy utility of the national COD data from six high-income countries with highly developed health information systems: Australia, Canada, Denmark, Germany, Japan and Switzerland.

B . 研究方法

Medical certification of death is a requirement in all the countries included in this study. Certifiers, mostly physicians and special health care

providers, are asked to complete a medical death certificate indicating what was, in their opinion, the sequence of morbid events leading to death. Subsequently, the information provided on the death certificate is coded by trained coders, applying the International Statistical Classification of Diseases and Related Health Problems (ICD) and its rules for COD coding, and compiled by health or statistical authorities (WHO 2016).

National data on population and COD coded to ICD-10, by age group and sex, were provided by Australia, Canada and Germany for 2016 and downloaded from the World Health Organization (WHO) website for Denmark, Japan and Switzerland for 2015 (WHO 2019). To evaluate the data sets, the ANACONDA tool (see Annex 1, Mikkelsen et al. 2020) that assesses the quality of national and subnational COD data was used.

The accuracy of the COD output is dependent on physicians providing enough information on the death certificate for coders to select and code the underlying COD. When this process is not completed correctly, the COD output may contain codes labelled “garbage” codes (Murray and Lopez 1996) because they are of little or no use for policy decision-making. Historically, “garbage” codes are defined as causes that cannot or should not be an underlying cause of death (Naghavi et al. 2010). The term, despite its inelegance, has now been an integral part of the literature for more than a quarter of a century.

Since ANACONDA uses the same concept and definition for these codes, we have retained the original terminology in the paper.

To provide additional insight into the provenance and policy implications of these codes, ANACONDA classifies garbage codes into two distinct typologies. In the first typology, garbage codes are grouped into five categories based on ICD concepts:

- Category 1: Codes relating to symptoms, signs and ill-defined conditions (most drawn from ICD Chapter XVIII); e.g. R99 Other ill-defined and unspecified causes of mortality).
- Category 2: Codes that are not valid as an underlying cause of death (e.g. T12 Fracture of lower limb).
- Category 3: Codes that represent intermediate causes of death (e.g. I50 Heart failure).
- Category 4: Codes that represent immediate causes of death (e.g. I46 Cardiac arrest).
- Category 5: Codes that represent insufficiently specified causes within ICD chapters or within a larger disease category (e.g. D48.9 Neoplasm of uncertain or unknown behaviour, unspecified).

ANACONDA also includes a second typology, which focuses much more on the potential impact that garbage codes might have on misguiding policy and planning (Naghavi 2020). In this typology, garbage codes are grouped into four impact levels, from “very high” (level 1) to “low” (level 4):

- Very high (level 1): This highest level represents causes for which the true underlying cause could be a communicable or non-communicable disease, or the result of an injury (e.g. septicaemia).
- High (level 2): These are causes with substantial negative impact, but where the true cause is mostly limited to one of the three broad cause groups mentioned above, e.g. essential (primary) hypertension that can be due to different non-communicable diseases.
- Medium (level 3): CODs classified to the third level are only considered to have a medium negative impact for policy since, in this case, the underlying cause is likely to be within the same ICD chapter (e.g. unspecified cancer).
- Low (level 4): Causes classified as having low negative impact are those where the true underlying cause is likely to be confined to a single disease or injury group, such as unspecified stroke or unspecified pneumonia.

These “impact-level” categories can be further grouped into those that provide no or little useful information about the true underlying cause (levels 1–3), which we therefore refer to as “unusable”, and those in level 4 that provide sufficient information to guide public health interventions but not for research and technology development (Naghavi et al. 2010). We refer to the latter as “insufficiently specified” causes as they impair evidence-based health policy processes only to a limited extent. However, correcting these becomes increasingly important if our health information systems are to

appropriately guide research, hospital financial flows, resource allocation and healthcare strategies (WHO 2019). This is likely to be particularly of relevance in countries with ageing populations where most deaths happen in hospitals, primarily from non-communicable diseases.

In countries where unusable codes are assigned to a large proportion of all deaths, the true COD distribution can be seriously distorted and thereby mislead policy dialogue. This is particularly serious when garbage codes are common among the leading causes of death. From the input data, ANACONDA automatically provides a listing of the top-20 COD for males and females and indicates those that are considered to be unusable (levels 1–3) or insufficiently specified (level 4). The higher the number of these codes and the higher their ranking, the greater their impact on misinforming policy is going to be.

The relationship between age and garbage codes is also investigated with the ANACONDA tool to verify whether they are particular to certain age groups. Furthermore, given that some differences might exist in population age structure between the six countries, we used the global proportion of deaths by age from the latest Global Burden of Disease Study as the standard (Murray et al. 2018) to age-standardize the garbage codes in the countries.

Data completeness, a key indicator of data quality, was not considered, given that all six countries have civil registration systems that register all deaths. The focus of our data quality analysis therefore was limited to the levels, patterns and distribution of garbage codes.

C . 研究結果

Despite the six countries being from three different geographic regions—Europe, Asia-Pacific and North America—their health systems and socio-economic indicators are comparable (Annex 2 Table 1S). Life expectancy varies from 78 to 81 years for males and 83 to 87 years for females, with all having very low child mortality rates of 2–5 per 1000 live born. The total fertility rates and proportion of 65 years and above indicate that Australia and Canada have somewhat younger populations than Denmark, Germany, Japan and Switzerland. Switzerland, with a private health insurance system, spends significantly more money on health care per person than the other countries. The Socio-Demographic Index (SDI) (Wang et al. 2016), a measure of national development based on income, education and fertility, is high for all countries, especially Denmark, while Germany and Denmark are doing slightly less well than the others on the Health Access and Quality Index (HAQ). The VSPI(Q), a measure of the overall quality of mortality data calculated by ANACONDA, is the highest in Australia and the lowest in Japan.

In the six countries studied, the average proportion of unusable codes (levels 1–3) was 18%, being slightly lower (14%) in Australia and Canada, while higher in Japan, where one in four deaths is assigned an unusable cause (Table 1). Insufficiently specified codes (level 4), in addition, averaged 8%, varying from 6% in Switzerland to 11% in Japan. Three of the most common CODs in the insufficiently specified group are pneumonia, stroke and diabetes all unspecified. For these CODs, the certifier could have increased the utility of the information provided on the medical certificate of death by specifying whether the pneumonia was bacterial or viral, the stroke ischaemic or haemorrhagic and the diabetes type 1 or 2.

Given the highly developed status of the six countries, the distribution of deaths on the three broad GBD groups of health conditions as expected showed that communicable and maternal diseases as well as injuries are minor contributors to their disease burden (Table 1). Non-communicable diseases on average accounted for 67.8% of all causes of death. Only Japan showed an unlikely low proportion (58.5%) that points to the impact that garbage codes can have on the cause pattern of mortality.

Since most deaths in these countries occur at older ages, it might be expected that these age groups also account for most of the garbage codes. That is indeed the case; in all six countries, between 85 and 92% of the garbage codes occur at ages 65

years and over. However, garbage codes are not limited to the oldest ages; they also comprise a sizeable proportion of deaths in several other age groups, particularly in the younger adult age groups where they constitute between 20 and 30% of all deaths. Even for child deaths, we do not know the true underlying cause in 10% of cases (Fig. 1). Deaths at these ages are often entirely preventable, but to do so public policy must be guided by accurate and specific COD data and how they are changing.

The first ANACONDA typology classifies the total amount of garbage codes according to five categories of certification errors and shows the percentage of each in relation to total deaths, and as a percentage of the total number of garbage codes (Table 2). In all countries except Germany, insufficiently specified COD (Category 5) was the most common error. The reporting of intermediary instead of underlying COD (Category 3) was the second most frequent issue in Australia, Canada, Japan and Switzerland. In Denmark, the second most frequent reporting flaw was Category 1 (the reporting of signs, symptoms or other ill-defined COD), followed by the reporting of an intermediary COD (Category 3). It is to be expected that all countries assign some deaths to ICD-10 code R99 (Other ill-defined and unspecified causes of mortality) since there always will be deaths for which the cause was unknown. However, Denmark and Japan stand out by coding more than 7% of all deaths to

category 1, which is much higher than in any of the other countries.

On average, certifiers in the six countries reported an intermediary COD (Category 3) as being the underlying COD for 9% of deaths. This error was particularly common in Japan (13%) and Germany (11%), almost twice as high as in the other countries. Further investigation showed that “Heart failure, unspecified” (I50.9) and “Congestive heart failure” (I50.0) were used more frequently in Japan and Germany than in other countries. Regarding the two remaining categories, irrespective of country, very few doctors certified an impossible COD (Category 2) or just provided the immediate COD (Category 4). The second typology of garbage codes provides important insight into the potential impact that garbage codes might have in guiding or misguiding public policy. This categorization showed a similar pattern for all countries with the “very high” impact category being the biggest problem for all six countries, followed by the “low” impact category, except for Australia, where the order of these top two impact categories was inverted in comparison with all other countries (Table 3). However, of note, the percentage at the “very high” impact level showed substantial differences between countries. For instance, in Japan, 21% of all deaths and 58% of all garbage codes had a very high impact for policy, while in Australia the comparable figures were only 8% and 35%, respectively. A closer investigation of the specific codes revealed that

the three ICD codes that account for most of the “very high” impact garbage codes were “Other ill-defined and unspecified deaths” (R99), “Heart failure” (I50.9) and “Senility” (R54).

The high and medium levels typically only accounted each for 2–4% of all deaths in all countries. The most common misdiagnosis for the high level was Essential (primary) hypertension (I10), Unspecified external factor (X59) and Gastrointestinal bleeding (K92.2), while for the medium level it was Unspecified cancer (C80.9) for all. The low-impact garbage codes were generally between 6% (Switzerland) and 11% (Japan) of all deaths and, on average, were used for 8% of all deaths, accounting for 31% of all garbage codes. In the low-impact group, the biggest contributor was “Stroke not specified” (I64), except for Japan where certifiers seem to better distinguish between haemorrhagic or infarction stroke, but not between the different types of pneumonia).

Standardizing the age structure for each country had a minor impact on the proportions of garbage codes for five of the countries. Only in Japan, where a higher proportion of deaths occur at the very oldest ages, age standardization reduced the amount of total garbage codes by 6%. However, Japan still remained the country with the highest proportion of deaths assigned to garbage codes (Table 3). The unusable codes among the leading causes of death are identified by red cells in Table 4. All

countries, except Australia, had at least one cell with unusable codes (red) among the top 10 causes of male deaths and all had one or more among the top 20 causes (Table 4). Japan had four red cells, and three of these were among the top 10 causes. Denmark and Canada had one red cell, while Switzerland and Germany had two red cells in the top half of the ranking. The specific unusable causes were very similar across the countries and included Other ill-defined and unspecified deaths (R99), Unspecified heart failure (I50.9), Congestive heart failure (I50.0), Unspecified cardiac arrest (I46.9), Unspecified malignant neoplasm (C80.9), Senility (R54), Pneumonitis (J69) and Unattended deaths (R98). With the exception of Japan, all countries also had two cells with insufficiently specified causes (orange) among the 20 leading causes of death.

For females, the 20 top disease rankings were even more saturated with unusable and poorly specified disease groups and included, apart from those mentioned for males, Essential (primary) hypertension (I10), Septicaemia (A41.9) and Malignant neoplasm of overlapping lesion of bronchus and lung (C34.8). Japan and Switzerland each had five red cells among the leading causes of female deaths, Denmark four, Canada and Germany three and Australia two.

Fortunately, a relatively small number of ICD codes are responsible for the major share of the garbage codes in these countries. In Table 5, the most common garbage codes for each country

have been identified. If these relatively few codes were not used, it would lead to a 25% reduction in the total amount of garbage codes. In Japan, this could be achieved very easily by avoiding the use of two codes: Senility (R54) and Unspecified heart failure (I50.9). As shown in Table 5, Denmark, Germany and Switzerland would need to focus on three codes and Canada and Australia, respectively, on six and seven. In other words, significant reductions in garbage codes could be achieved if certifiers, instead of just certifying that patients died from old age, heart failure, hypertension, septicaemia and unspecified cancer, could more accurately report the sequence of events leading to death, including the underlying cause of that sequence. As noted above, the problem with the code R99 (ill-defined and unspecified causes of mortality) is not that it cannot be used but that it is over-used, e.g. it is unlikely that no cause could be identified in Denmark for 5% of all deaths.

D. 結論

Although the countries included in the study have highly developed mortality information systems and have been producing COD data aligned with international standards for many years, the assessment of their data still revealed that there were some unexpected deficiencies in their statistics that could have significant implications for policy dialogue, monitoring health progress and evaluating intervention impact. This would appear to be due, in large part, to the lack of standardized instructions for medical certifiers

about how to correctly complete the death certificate. All six countries declared that such basic information is not part of the standard training provided to young doctors. While physicians may not need to be trained to use the ICD, they should at least be taught how to properly certify the sequence of events leading to death to ensure that coders can identify correctly the underlying cause that led to the person's death. It is this information that is critical for guiding public health policies to further reduce premature mortality and address the rising costs of health care.

In systems where all deaths are registered with a cause, the bias in the data is largely determined by the level and type of garbage codes they contain. To certify that 17% of all deaths in the 70-plus age group were due to old age (Japan) is unhelpful if health authorities want to have a better understanding of disease management in later life. Most people in the considered countries die at older ages and are likely to have had frequent contact with the health system. It is reasonable therefore to assume that comprehensive medical records exist that should allow physicians to more accurately certify deaths. The tendency to assign "old age" as a COD strongly suggests that the certifier has not been trained and is unaware of the important public health use of the death certificate.

While all countries will have a small proportion of deaths where the circumstances leading to death are either

not known or cannot be further specified, hence justifying the use of the R99 code “Other ill-defined and unspecified deaths”, it is of concern when this cause appears among the leading causes. The same can be said for “Unspecified heart failure”, which is an intermediary COD and which accounts for between 3 and 14% of the garbage codes in the six countries examined. Rather, physicians should specify the underlying cause that led to death, which in the case of heart failure could, among others, be myocardial infarction, chronic renal failure, cerebrovascular accident, poisoning and haemorrhage.

Decreasing the proportion of deaths that are of no or little policy value should be a priority for all country health information systems. This assessment and analysis have shown that it is possible for health authorities to quickly obtain insight into the quality problems and certification errors in the data, which can assist them to take corrective action. The results of the analysis reveal the main certification errors committed by doctors and identify the CODs that introduce bias into the leading causes of death. As demonstrated, the quantity, severity and type of garbage codes appearing in the national cause of death data vary across countries, suggesting that certification problems and coding may be somewhat culture specific. Identifying the specific garbage codes that produce the most unusable data in each country is an essential first step so that strategies can be tailored and developed to determine effective ways to train or inform certifiers. This

will help certifiers to be more specific when they complete medical death certificates and to avoid committing errors that lead to garbage codes. Being aware of quality problems in the data can also assist the authorities responsible for publishing the COD data to provide the explanations needed to correctly interpret the statistics. For instance, in Australia the government agency that publishes the COD data is careful to add the poisoning agent when it publishes data on the accidental poisoning garbage codes such as X42 and X44, thus increasing the information content of the data for policy that normally would be missing from these garbage codes.

National COD data represent a compilation of data from different geographic areas and health facilities, which may vary in their death certification practices, and hence accuracy. It is therefore advisable that countries undertake subnational assessments to verify how certification practices vary and tailor intervention strategies accordingly with more local approaches. For instance, lack of certain diagnostic imaging and analysis and under-staffing can lead to less-than-optimal medical records and make it even more challenging to correctly certify the COD.

Undertaking this type of assessment and communicating the findings to health authorities and medical associations will result in greater

awareness of the need to pay more attention to the quality of medical certification. Medical schools in all six countries should give higher priority to the certification duty of their profession, and correct medical certification certainly should be included as a compulsory element of the induction programs for interns. Doctors perform a very important public health function in documenting the COD of their patients, not only for the families of the deceased, but also for society. They are generally unaware of this, or of its importance for public policy. Without reliable and detailed information on the leading CODs, and how they are changing, health planning and policy will be less cost-effective than otherwise might be the case, potentially resulting in lost opportunities to improve population health.

While ANACONDA cannot verify whether the physician diagnosed the correct COD or whether a COD was miscoded, it can detect whether the code assigned is a valid underlying cause and whether the certifier originally completed the death certificate according to ICD guidance. Adopting the ICD classification without adhering to its rules and standards for coding and guidance for certification will not provide good-quality COD information for public health use. This study has demonstrated that even for countries with very advanced health information systems, it is very informative to undertake an assessment of the COD data as the information content may be reduced by a high proportion of unusable and insufficiently specified causes.

Once the data have been evaluated and the specific problems revealed, focused action should be taken to reduce the amount of garbage codes by, for example, introducing certification training for hospital interns and awareness raising in the medical community of the important functions of the death certificate for the national health information system. The large and complex bureaucracy that all countries have established to collect these data needs to meet the demands of increasingly sophisticated and complex health systems and provide them with the detailed information required for avoiding premature deaths and keeping people alive and healthy for as long as possible.

* 本稿は、「Mikkelsen L, Iburg, KM, Adair T, Fürst T, Hegnauer M, von der Lippe E, Moran L, Nomura S, Sakamoto H, Shibuya K, Wengler A, Willbond S, Wood P, Lopez AD. **Assessing the quality of cause of death data in six high-income countries: Australia, Canada, Denmark, Germany, Japan and Switzerland.** Int J Public Health. 2020 Jan; 65, 17–28. 」に掲載された。

E. 研究発表

1. 論文発表

Mikkelsen L, Iburg, KM, Adair T, Fürst T, Hegnauer M, von der Lippe E, Moran L, Nomura S, Sakamoto H, Shibuya K, Wengler A, Willbond S, Wood P, Lopez AD. **Assessing the quality of cause of death data in six high-income countries: Australia, Canada, Denmark,**

Germany, Japan and Switzerland. Int J Public
Health. 2020 Jan; 65, 17–28.

2. 学会発表

特になし

F. 知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得

特になし

2. 実用新案登録

特になし

3. その他

特になし

