

EVIDENCE BASED PUBLIC HEALTH POLICY AND PRACTICE

Evaluation of the quality of cause of death statistics in rural China using verbal autopsies

Lijun Wang, Gonghuan Yang, Jiemin Ma, Chalapati Rao, Xia Wan, G Dubrovsky, Alan D Lopez

J Epidemiol Community Health 2007;61:519–526. doi: 10.1136/jech.2005.043927

Background: Death registration systems in rural China are in a developmental stage. The Disease Surveillance Points (DSP) system provides the only nationally representative information on causes of death. In this system, there are no standard procedures or instruments for ascertaining causes of death; hence available statistics require careful evaluation before use.

Aim: To assess the reliability of data from the DSP.

Methods: 14 DSP sites were selected through stratified sampling, enrolling 2482 deaths registered during June–November 2002. Defined verbal autopsy (VA) procedures were used to derive underlying causes of death. κ Measures of agreement between VA and registered diagnoses were computed. VA diagnoses were used as references to compute sensitivity and positive predictive values. Finally, patterns of cause-specific discordance between the two data sources were studied.

Results: κ Scores indicate only moderate agreement for chronic obstructive pulmonary disease (COPD) and ischaemic heart disease (IHD). The results also suggest that COPD is undercounted in registration data. A significant degree of cross attribution of cause was found between the two data sources. Overall, the VA was successful in limiting deaths with ill-defined causes to <10%.

Conclusions: Diagnoses based on structured symptom questionnaires in the VA approach seem plausible and reliable as compared with lay-reported diagnoses in registration data. Concerns with attribution of cause of death due to COPD, IHD and tuberculosis in registration data suggest caution in their use for research and health programme purposes. The VA methods tested in this study offer promise for implementation in the routine registration system.

See end of article for authors' affiliations

Correspondence to:
A D Lopez, School of Population Health, University of Queensland, Public Health Building, Herston Road, Herston, Brisbane, Queensland 4006, Australia; a.lopez@sph.uq.edu.au

Accepted 4 January 2006

Reliable information on the levels and causes of mortality is critical for the assessment of population health status. Although complete vital registration systems with medical certification of the cause of death are the "gold" standard source for such data, they are difficult to establish and expensive to maintain in large populations such as that of China. Medical certification of the cause of death is dependent on the availability of a doctor at death. In rural China, most deaths occur at home, in the absence of medical attention and with limited contact with health services during the illness preceding death. Given this background, statistics on registered causes of death in rural China require careful evaluation if they are to be useful for epidemiological research and public health planning.

Mortality registration and cause of death reporting systems in China are in a developmental stage. A detailed description of the systems in place for the registration and compilation of cause of death statistics in China is available elsewhere.¹ In brief, although close to a billion people reside in rural China, cause of death statistics are compiled for only 0.6% of this population through the nationally representative Disease Surveillance Points (DSP) system. This system operates in 78 townships (Xiangs) in rural areas, with populations ranging from 30 000 to 100 000 in each site. There are no standard procedures or instruments for the collection and reporting of causes of death. A total of about 35 000 deaths are registered annually in these rural sites.^{2–3} Causes of death are investigated and registered through the Disease Prevention Unit in the local township hospitals, on the basis of unstructured interviews of the family of the deceased. In this paper, we report on the findings of an evaluation of the quality of cause of death attribution in rural China on the basis of a sample of 2482 deaths in 14 DSP sites. Clearly, understanding the reliability of

cause of death data for one sixth of the world's population has important implications for global burden of disease assessments as well as for health sector planning in China.

Evaluation framework

A critical element in any evaluation of data quality is the assessment of content validity of the reported causes of death, by comparing them with gold standard reference diagnoses for the same deaths,⁴ preferably for a representative sample of deaths occurring in the population. In rural China, few deaths occur in a health facility, which excludes the possibility of deriving reference diagnoses for most deaths in any representative sample. Hence, it is not possible to conduct a classical validation study in rural China. The next best option is to assess the reliability of cause of death reporting in the routine registration system by comparing registered diagnoses with diagnoses derived through an independent, standardised cause-of-death ascertainment strategy. In this study, we used a standardised verbal autopsy (VA) procedure for the cause-of-death ascertainment concurrently implemented and validated in urban sites in China,⁵ to derive independent diagnoses for each death in the study sample.

In principle, the VA method is more reliable than the procedures used in the routine registration system since a structured, detailed set of questions about symptoms should lead to better cause of death certification and coding of the underlying cause of death. The validation study in urban areas of China used detailed hospital records to derive reference

Abbreviations: CAMS, Chinese Academy of Medical Sciences; COPD, chronic obstructive pulmonary disease; DSP, Disease Surveillance Points; ICD, international classification of disease; IHD, ischaemic heart disease; TB, tuberculosis; VA, verbal autopsy

diagnoses. Although admittedly a blunt method, the VA was found to have good validity for stroke, cancers and perinatal conditions, and although less satisfactory for other leading causes such as chronic obstructive pulmonary disease (COPD), ischaemic heart disease (IHD) and tuberculosis (TB), the misclassification by VA was compensatory, resulting in minimal distortion of the overall cause-specific mortality patterns in the study sample. With this knowledge of its intrinsic operational characteristics from the urban validation study, we have applied the VA method as a proxy gold standard to estimate the reliability of the routine registration system in rural areas. Although formal validation is not possible in rural China, it is likely that cause of death patterns based on the uniform application of a symptom-based questionnaire will be more reliable than those based on informal enquiry of families about their opinion as to the cause of death of family members.

For those causes for which the VA is likely to predict correctly the underlying cause of death, the results of the VA re-investigation might be used to correct the cause of death pattern reported by the routine registration system, for which statistical procedures have been proposed.⁶ Moreover, the re-investigation is likely to reduce substantially the proportion of deaths coded to ill-defined conditions and thus can be directly used in burden of disease estimation exercises, as has been done in India⁷ and Thailand,⁸ or for epidemiological research.⁹

METHODS

The DSP system operates in a stratified random sample of 78 counties across four rural socioeconomic strata, ranging from rural 1 (best off) to rural 4 (worst off). For this project, 14 sites were randomly selected, including two from rural 1, and four each from the other three strata (fig 1). The poorer rural strata were better represented in the study since it was anticipated that the quality of cause of death data would be more questionable in these counties.

The reference period for the study was 1 June 2002 to 30 November 2002. All deaths registered in each project site during this period were recruited into the study. In each site, the local

Chinese Centre for Disease Control field office (responsible for the routine compilation of data) submitted a complete list of deaths registered during the reference period to the Project Office at the Chinese Academy of Medical Sciences (CAMS) in Beijing. The list, arranged in chronological order of date of death, included the name and address of the deceased, informant details, place of death and the registered cause-of-death. This resulted in a total of 2482 cases across the four strata (table 1). The Project Office subsequently abstracted a list omitting the cause of death as the set of cases for field investigation.

For each selected site, a medical doctor from the corresponding township hospital was trained in the use of standard VA field instruments, and in VA interviewing skills at a central training programme conducted in Beijing. Each trained investigator was provided with the list of cases for field investigation, and the interviews were conducted during January–December 2003. Senior research management staff from the CAMS visited all field sites twice to provide onsite refresher training, supervision and feedback on interviewing skills and data recording. For the two sites in western China (Hetian and Menyuan), the study reference period was extended to March 2003 and the data collection to March 2004, owing to smaller population size and greater under-reporting of deaths. The interval between death and household interviews ranged between 3 months and 1 year in most cases. After household interviews, each completed VA instrument was submitted to the CAMS for further processing. First, a panel of three medical practitioners (internists) from the Military Hospital in Beijing with similar training reviewed the VA responses, and assigned a cause of death for each case on the basis of the standard international form of medical certificate of cause of death.¹⁰ Trained medical coders at the World Health Organization international classification of diseases (ICD) Collaborating Centre in Beijing then selected and coded the underlying cause of death from each certificate, according to the rules and specifications of the *ICD and health related problems*, 10th revision (ICD-10).¹⁰ In comparison, doctors at township



Figure 1 Project sites from the Disease Surveillance Points System in rural China.

Table 1 Distribution of study sample by rural strata according to leading causes reported by the routine Disease Surveillance Points system

Cause	ICD codes	Rural 1	Rural 2	Rural 3	Rural 4
Cerebrovascular disease	I60–I69	62	185	180	55
COPD	J40–J47	21	124	98	9
Other cardiac diseases	I26–I51	31	31	38	124
Ischaemic heart disease	I20–I25	12	73	70	14
Ill-defined conditions	R00–R99	34	15	28	60
Liver cancer	C22	22	39	21	10
Stomach cancer	C16	10	48	27	6
Lung cancer	C33–C34	17	35	31	5
Road traffic accidents	V01–V99	13	27	36	5
Suicide	X60–X84	11	24	22	7
Hypertensive diseases	I10–I15	11	15	23	10
Drowning	W65–W74	4	15	9	20
Oesophageal cancer	C15	6	16	22	1
Liver diseases	K70–K77	5	11	17	10
Perinatal conditions	P00–P96	2	4	12	18
Tuberculosis	A15–A19	3	10	9	12
Pneumonia	J12–J18	1	5	5	21
All other cancers	*	18	58	29	10
All other external causes	W00–W64, W75–X59, X85–Y89	9	49	26	24
All other causes	All remaining codes	33	72	116	61
Total deaths	–	325	856	819	482

COPD, chronic obstructive pulmonary disease; ICD, international classification of diseases.
*C00–C14, C17–C21, C23–C32, C35–D48.

hospitals certify causes of death in the routine registration data, and underlying causes are selected and coded by a team at the CAMS. Summary tabulations of deaths by age, sex and underlying cause were developed for each data source (registration and VA investigation), aggregated by cause according to the Mortality Tabulation List 1.¹¹ Diagnoses from registration data were compared with VA diagnoses on the basis of the κ statistic of chance corrected agreement between the two data sources for the leading causes of death in the study sample.¹² Using the VA diagnoses as a proxy gold standard, the inferred “sensitivity” and “positive predictive values” of the registration data for each cause were computed. Under the same assumption, a comparative analysis of cause-specific mortality proportions for each cause category from the two data sources was used to determine the extent to which the routine registration system over- or under-reported specific causes. Finally, matrices were constructed to describe details of the cause-specific disagreement between the two data sources. Since the study sample is a close approximation of the registered cause of death profile in rural China (fig 2), the findings of such descriptive analyses provide evidence for correcting routine registration data in rural areas and estimating overall cause-specific mortality in rural China.

RESULTS

Table 1 shows the distribution of the study sample across the four socioeconomic strata according to leading causes as suggested by the routine registration system in these project sites. It would appear that cause attribution by the routine system in the worst-off socioeconomic stratum (rural 4) is problematic, with a high proportion of deaths assigned to “other cardiac diseases” (25.7%) and “ill-defined conditions” (12.5%), compared with other strata. On the other hand, the higher proportions of deaths due to perinatal causes, drowning, TB and pneumonia in this stratum are not inconsistent with lower socioeconomic status. Overall, a comparison of the proportionate distributions by cause from the aggregated study sample and the latest published DSP data from rural areas during 1997–9 (fig 2) suggests that the study sample is reasonably representative of the overall DSP system in rural

areas, the most notable exceptions being the lower proportionate representation of deaths due to COPD and pneumonia in the study sites.

Agreement scores

Table 2 gives the results of comparative analyses between the registration diagnoses and the VA-based re-investigation of causes-of-death in the study sample. κ measures of chance corrected agreement >0.4 represent fair to excellent agreement according to the rating scale proposed by Landis and Koch,¹³ and are observed for almost all the leading causes-of-death in the table. Moderate or poor agreement ($\kappa < 0.4$) was found only for TB, hypertensive diseases and poisonings, among the classified causes. Interestingly, κ was only marginally greater than 0.4 for COPD ($\kappa = 0.41$) and IHD ($\kappa = 0.42$) suggesting significant concerns with the coding of these conditions in rural China. The low κ values observed for “ill-defined causes” and “other cardiac diseases” are reassuring since one would expect the VA procedures to be more likely to assign classifiable cause-of-death diagnoses for these cases.

Cause-specific mortality fractions from the registered data and the VA procedures differ by $>25\%$ only for falls, TB and poisonings. This implies that if the VA diagnoses are a reasonable reflection of the true cause of death pattern, then the routine registration system is probably performing well for some causes (eg, stomach cancer and diabetes), but is yielding cause-specific mortality fractions for many of the leading causes-of-death that are only within 25% of their true value. Although this may be a reasonable basis to ascertain broad policy priorities, it suggests caution if the data are to be used for epidemiological end points.^{14 15}

Misclassification analysis

Table 3 shows the major patterns of disease misclassification, at least when compared with the VA-based diagnoses. Importantly, although the number of deaths due to stroke in the routine system (482) is similar to that suggested by the verbal autopsy enquiry (497), only 366 or 75% of the VA-based deaths due to stroke were identified as such by the routine system. Significant numbers of deaths due to stroke were coded to COPD (33), other cardiac diseases (18), IHD (17) or

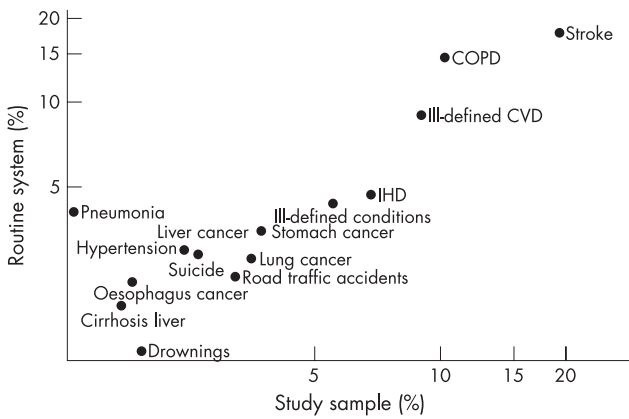


Figure 2 Correlation between proportionate distribution by cause in the study sample and Disease Surveillance Points registration system in rural China, 2002 (graphed in logarithmic scale). COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; IHD, ischaemic heart disease.

hypertensive disease (15). Thus, although most of the misclassified deaths due to stroke fall within the broad category of cardiovascular diseases, the implications for research or planning that requires more specific epidemiological detail are substantial. Interestingly, the table suggests that COPD may be undercounted as a cause of death in rural China, with the routine system identifying 224 or 75% of probable cases (297), as identified by VA. On the other hand, IHD does not seem to be undercounted in this sample of rural deaths and, if anything, may be 30% or so over-represented in rural mortality data.

A more stringent measure of reliability is the assessment of sensitivity of the registration system in accurately diagnosing causes of death compared with the cause suggested by the VA procedure. High sensitivity and positive predictive value scores

(>75%) are observed only for some external causes, cancers, diabetes, cerebrovascular disease and perinatal conditions. Consistent with the findings from the κ analyses, leading causes of death such as IHD and COPD have low “validity” scores against this standard.

One of the most striking observations from table 3 is the poor agreement between the two data sources for deaths coded to “other cardiac diseases” (ICD codes I26–I51), with only 14 of the 224 (6.2%) events reported in the registration data being confirmed by the VA. Rather, the VA classified most of these deaths (45.5%) to COPD. Further investigation revealed that for these deaths, the cause of death registered in the routine system was labelled as “pulmonary heart disease”, the code for which is I27.9 (which falls in the “other cardiac diseases” cause grouping), whereas the VA identified the cause as “heart disease secondary to chronic obstructive pulmonary disease”, which under the ICD rules is coded to J44.9, the code for COPD.¹⁰ This difference in semantics seems to be responsible for the generalised under-reporting of COPD by the routine registration system in rural areas. This has important implications for the rank order and magnitude of COPD as a cause of death in rural China.

Another condition for which agreement is low is TB (44%). Although the disagreement between TB and COPD seems compensatory, 11 deaths classified by the routine system as due to “other cardiac diseases” were diagnosed as TB by the VA. This suggests a possible under-reporting of TB by the registration system of almost 50% in this study sample, although caution is required in making population level inferences because the number of deaths due to TB is small.

Disagreement between the two data sources for other leading cardiovascular, respiratory and neoplastic causes of death seems to be largely compensatory, resulting in minimal changes in the overall cause-specific mortality profile. Interestingly, the proportion of ill-defined diagnoses suggested by the VA (9%) is almost twice that in the routine system, (5.5% (137/2482)), and is a more plausible fraction given the procedures involved.

Table 2 Measures of agreement between registration data and verbal autopsy diagnoses, and implied validity* of registered mortality fractions by cause in the Disease Surveillance Points system in rural China, 2002

Category	Deaths (n)		κ	Sensitivity*	PPV*	Change in cause-specific mortality fraction (%)
	In registration data	From VA re-investigation				
Drowning	48	44	0.91	95.5	87.5	8.3
Road traffic accidents	81	85	0.86	84.7	88.9	-4.9
Liver cancer	92	86	0.79	82.6	77.2	6.5
Diabetes	27	29	0.78	75.9	81.5	-7.4
Oesophageal cancer	45	52	0.76	71.2	82.2	-15.6
Stomach cancer	91	89	0.73	75.3	73.6	2.2
Suicide	64	63	0.73	74.6	73.4	1.6
Perinatal conditions	36	36	0.72	72.2	72.2	0.0
Colorectal cancers	29	24	0.71	79.2	65.5	17.2
Cerebrovascular diseases	482	497	0.69	73.6	75.9	-3.1
Lung cancer	88	79	0.68	73.4	65.9	10.2
Liver disease	43	54	0.57	51.9	65.1	-25.6
Falls	28	44	0.52	43.2	67.9	-57.1
Other external causes	55	51	0.46	49.0	45.5	7.3
Pneumonia	32	40	0.44	40.0	50.0	-25.0
Ischaemic heart disease	169	128	0.42	53.1	40.2	24.3
COPD	252	297	0.41	44.1	52.0	-17.9
Tuberculosis	34	50	0.35	30.0	44.1	-47.1
Hypertensive diseases	59	50	0.33	38.0	32.2	15.3
Poisoning	14	25	0.30	24.0	42.9	-78.6
Ill-defined conditions	137	229	0.28	26.64	44.53	-67.2
Other cardiac disease	224	73	0.05	19.2	6.3	67.4
All other causes	352	357				-1.4
Total deaths	2482	2482				

COPD, chronic obstructive pulmonary disease; PPV, positive predictive value; VA, verbal autopsy. A positive change in cause-specific mortality fractions suggests over-reporting by the routine registration system, a negative change suggests under-reporting. *To measure validity, the VA diagnoses were used as “proxy” reference diagnoses, hence the use of the qualification term “implied”.

Table 3 Cross tabulation of registration and verbal autopsy diagnoses for specific communicable and non-communicable diseases in rural China, 2002

VA diagnosis	Registration diagnoses																All other diseases	Total					
	Stroke	COPD	Other cardiac diseases	IHD	IHD conditions	Liver cancer	Liver	Stomach cancer	Lung cancer	Hypertensive disease	Oesophageal cancer	Liver diseases	Perinatal conditions	Tuberculosis	Pneumonia	Other digestive disorders			Colorectal cancer	Diabetes	Rheumatic heart disease	Congenital malformations	Other cancers
Stroke	366	33	18	17	9	1	2	15						2	3	1	3	1	3	1		25	497
COPD	6	131	102	5	17	3	5	3			2			9	1	3	3	1	1	1		5	297
Other cardiac diseases	11	13	14	16	5	1	1	2		1									5			4	73
IHD	15	13	14	68	5			5						2	1	4	1	1	4	1		2	128
IHD-defined conditions	40	19	22	31	61	1		4		2	3			2	3	4	1	1				35	229
Liver cancer	1	1	2			71	1	3	2		2			1		1					3	0	86
Stomach cancer	6	4	4			2	67	4		2				2		1	1					1	89
Lung cancer	2	2	4			1	4	58	1	1				1					1		2	3	79
Hypertensive disease	6	4	4	4	7			19						1								4	50
Oesophageal cancer	2		2	1	1	1	4	4		37												1	52
Liver diseases	1			2	2	7		2			28					3		1				8	54
Perinatal conditions	1				4							26			3					1		1	36
Tuberculosis	8	11	2	2	2			3						15	1			1				7	50
Pneumonia	2	11	4									2		16	1					1		4	40
Other digestive disorders	1	3	3	8	1									1	1	12	2		1	1		13	49
Colorectal cancer	1					1	1	1								19						1	24
Diabetes	1	1	2	2	2			2										22	4			2	29
Rheumatic heart disease	2	2																				1	13
Congenital malformations												2									13	3	19
Other cancers	2	2	1	3		2	2	1	1					3	2	7	2				8	5	29
All other diseases	22	12	14	11	16	3	7	4	5	2	8	6	6	3	2	0	3	2	3	2	6	425	559
Total	482	252	224	169	137	92	91	88	59	45	43	36	36	34	32	31	29	27	22	20	19	550	2482

COPD, chronic obstructive pulmonary disease; IHD, ischaemic heart disease; VA, verbal autopsy.

Table 4 Cross tabulation between registration and verbal autopsy diagnoses for external causes in rural China, 2002

VA diagnosis	Registration diagnosis									Total
	RTA	Suicide	Other external causes	Drowning	Falls	Poisoning	Fires	Assault	Natural causes	
RTA	72		4	1	2		1		5	85
Suicide	3	47	2	1		2		1	7	63
Other external	3	1	25		1		4	1	16	51
Drowning				42					2	44
Falls			8	1	19				16	44
Poisoning	1	10	1			6			7	25
Fires							2		0	2
Assault	1		3		1			4	4	13
Natural causes	1	6	12	3	5	6	0	1	2121	2155
Total	81	64	55	48	28	14	7	7	2178	2482

RTA, road traffic accidents; VA, verbal autopsy.

Table 4 shows the principal differences in diagnoses between the two datasets for external causes of death. If the VA diagnoses were likely to be a more reasonable assessment of the true cause of death in the study sample, then the registration system seems to under-report deaths due to external causes by about 1%. This discrepancy mainly arises from the VA re-classifying 16 deaths reported as due to "natural causes" by the registration system as due to falls, and a similar number to external causes other than those specified in the table. Thus, if the VA causes are correct, the implication is that the registration system under-reports deaths due to falls by 57% (table 2). The VA also re-classified 10 deaths for which the registered cause was given as "suicide" to poisoning, which accounts for much of the substantial undercount of poisonings in the rural registration data. For the remaining external causes, the patterns of disagreement seem largely compensatory, and do not distort the rank order and magnitude of the reported cause-specific mortality.

DISCUSSION

A recent evaluation of cause of death statistics in China⁴; and their uses,¹⁶ suggested an urgent need to evaluate critically the content validity of available mortality data. We provide such an assessment of routine statistics from urban China in a companion paper.¹⁷ The availability of reference diagnoses (ie, from clinical case histories) for validation in urban areas makes such an assessment more meaningful. Yet, since the vast majority of deaths in China occur in rural areas,¹⁸ it is of obvious importance to assess the validity of rural mortality data using similar procedures. However, the lack of reliable, detailed clinical information on deaths in rural areas precludes such an assessment.

This study has attempted to evaluate the quality of cause-of-death data in rural China using standard VA procedures, arguably the next best option. Overall, the registered cause-specific mortality statistics seem to be reasonably reliable for several leading causes-of-death, but substantially incorrect for several other major conditions including COPD and TB, both of which are significantly under-reported in the routine statistics, and IHD, which seems to be significantly over-coded. The probable misclassification of COPD is particularly worrying given that recent global estimates of mortality by cause, which have utilised the Chinese data, suggest that COPD kills about one million people a year in China, or 50% of the global total.¹⁹ The discrepancy in diagnosing deaths due to COPD may well be an artefact of the methods for certification and coding of death certificates in the routine registration system, and suggests the need for better training of certifying doctors and coders in the correct ICD procedures.

Inferences about the true underlying pattern of causes-of-death in rural China from our study need to be drawn with extreme caution. In the absence of a true gold standard diagnosis based on autopsies, or a good approximation to a gold standard based on detailed clinical evidence, it is unclear how reliable VA procedures will be in identifying the true cause of death for comparative purposes. Validation of the VA in urban areas of China suggests that the instrument is only approximately valid for several leading causes and this is a very questionable basis to assess reliability. Rather, our findings should be interpreted more as an alternative set of diagnoses based on standardised procedures than an attempt at validation of the rural cause of death data. Even where the operational

What this paper adds

- In rural China, the Disease Surveillance Points (DSP) system is responsible for the collection, collation and reporting of information on causes of death, from a nationally representative sample of 78 surveillance points distributed throughout China. Data collection procedures and instruments differ between sites, and largely comprise of information gathered through household interviews of relatives of the deceased. This study was designed to test the implementation of a set of standard "verbal autopsy" procedures for collecting such data in a selected sample, and compare causes of death identified using these procedures with registered diagnoses for each death in the study sample.
- The study found that there was an overall acceptance of standard procedures for data collection across sites. Also, these procedures were successful in limiting the proportion of deaths classified to ill-defined causes to less than 10%, implying the feasibility of their wider implementation. Comparative analyses suggest an undercount of chronic obstructive pulmonary disease mortality in registration data, with significant degree of cross attribution between the two data sources, especially from ischaemic heart disease, and tuberculosis. These findings suggest caution in the use of registration data for epidemiological research and health programme purposes, and highlight the need for further studies to be conducted periodically to assess the reliability of cause of death attribution using verbal autopsies.

Policy implications

- This study has highlighted the need for standardised cause of death reporting mechanisms in the Disease Surveillance Points (DSP) system.
- The system requires strengthening in the form of increased medical input into the cause of death registration process.
- Capacity building is required in verbal autopsy interviewing skills, cause of death certification and ICD coding.
- Policies that enable the implementation of the above could result in more reliable cause of death statistics from the DSP in about 5 years.

characteristics of the VA procedures are known, and hence can be adjusted to suggest true cause-specific mortality fractions, the problem of transferability of validation characteristics from one (urban) population to another (rural) creates additional uncertainty about the reliability of the VA diagnoses as a comparator. Additionally, the small sample sizes for some diseases such as TB make it difficult to draw conclusions about possible misclassification patterns for the large rural population in China.

Having said this, the performance of the VA in limiting the proportion of deaths classified to ill-defined conditions to <10% is impressive, and consistent with conventional norms for cause of death statistics under this criterion.²⁰ Where the VA investigation reveals no significant evidence, the procedures recommend the assignment of a symptom-based or ill-defined cause, and the fact that this was relatively uncommon in rural China suggests that the approach may have promise as a cost-effective means of data collection for almost one billion people. However, more research is needed to ascertain the reliability of specific diagnoses; at present, there is no evidence to support or refute the diagnoses reported in the routine registration data, whether classified adequately or not.

Although the instruments and procedures for VA have only been formalised and termed as such for the past 20 years or so,^{21–23} similar methods have been used to derive causes of death in earlier times,²⁴ and widespread adoption of VA procedures could very well hasten the development of useable cause-of-death statistics in much of the developing world.

An evaluation of data quality such as in this study serves additional purposes. The use of formal procedures for cause-of-death ascertainment in the study enables one to test their applicability and acceptance in rural China. The methods could therefore be applied more widely to improve the overall quality of cause of death data in rural China. Although data from the registration system have already been widely used for descriptive analyses of burden of disease¹⁹ and for evaluating health policy,²⁵ the reliability and validity of the data are critical for these functions, as well as to monitor the health status of rural China, given the current differentials in access to healthcare²⁶ and socioeconomic status²⁷ in this vast country. In this situation, mortality indices, although crude, could be adequate to assess the magnitude of these differentials.^{28–29}

Also, having greater confidence in the cause of death data emanating from rural China is essential if Chinese health planning is to be more reliably informed and global disease

burden estimates are to be more confidently used. Although the methods used in this study do not allow formal validation of data from rural China, they are nevertheless suggestive of major misclassification errors for specific causes of scientific and public health interest. Certainly, epidemiological research based on these data that requires information on disease-specific end points should proceed with great caution, although cause-of-death patterns for broader cause groups are likely to be more reliable since misclassification is often internal to broad disease categories. The research does suggest significant changes to population-level cause-specific patterns being estimated from these data, particularly for COPD and TB, and, together with the findings from validation research in urban China, could be used to re-evaluate the pattern and stage of epidemiological transition in China.

ACKNOWLEDGEMENTS

We thank the US National Institute on Aging for providing financial support for this project (research grant number PO1 AG17625). The funding agency had no role in the study design and data collection, data analysis or in the preparation of this manuscript.

Authors' affiliations

Lijun Wang, National Centre for Disease Control and Prevention, Beijing, China

Gonghuan Yang, Jiemin Ma, Xia Wan, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, Beijing, China

Chalapati Rao, G Dubrovsky, Alan D Lopez, School of Population Health, University of Queensland, Herston, Brisbane, Queensland, Australia

Competing interests: None.

REFERENCES

- 1 **Yang G**, Hu J, Rao KQ, *et al*. Mortality registration and surveillance in China: history, current situation and challenges. *Popul Health Metr* 2005;**3**:3.
- 2 **Chinese Academy of Preventive Medicine**. *Annual Report on Chinese Disease Surveillance 1992*. Beijing: CAMP, 1994.
- 3 **Chinese Academy of Preventive Medicine**. *Annual Report on Chinese Disease Surveillance 1995*. Beijing: CAMP, 1996.
- 4 **Rao C**, Lopez AD, Yang GH, *et al*. Evaluating national cause of death statistics: principles and application to the case of China. *Bull World Health Organ* 2005;**83**:618–25.
- 5 **Yang G**, Rao C, Ma J, *et al*. Validation of verbal autopsy procedures for adult deaths in China. *Int J Epidemiol* 2005;**35**:741–8.
- 6 **Chandramohan D**, Setel P, Quigley M. Effect of misclassification of causes of death in verbal autopsy: can it be adjusted? *Int J Epidemiol* 2001;**30**:509–14.
- 7 **Mahapatra P**, Rao C. *Causes of death in rural areas of Andhra Pradesh, 1998*. Working Paper 40/2000. Hyderabad, India: Institute of Health Systems, 2000.
- 8 **Ministry of Public Health**. *Burden of disease and injury in Thailand: priority setting for policy*. Bangkok, Thailand: Ministry of Public Health, 2002.
- 9 **Gajalakshmi V**, Peto R, Kanaka S, *et al*. Verbal autopsy of 48 000 adult deaths attributable to medical causes in Chennai (formerly Madras), India. *BMC Public Health* 2002;**2**:7.
- 10 **World Health Organization**. *Mortality: guidelines for certification and rules for coding. International statistical classification of diseases and health related problems—10th revision. Vol 2: instruction manual*. Geneva: World Health Organization, 1993:30–65.
- 11 **World Health Organization**. *Special tabulation lists for mortality and morbidity; mortality tabulation list 1. International statistical classification of diseases and health related problems—10th revision. Vol 1*. Geneva: World Health Organization, 1993:1207–10.
- 12 **Kirkwood B**, Sterne J. Measurement error: assessment and implications. *Essential medical statistics*. Massachusetts: Blackwell Science, 2003:429–39.
- 13 **Landis JR**, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74.
- 14 **Lopez AD**. Counting the dead in China. Measuring tobacco's impact in the developing world. *BMJ* 1998;**317**:1399–400.
- 15 **Niu SR**, Yang GH, Chen ZM, *et al*. Emerging tobacco hazards in China: 2. Early mortality results from a prospective study. *BMJ* 1998;**317**:1423–4.
- 16 **Phillips MR**, Li X, Zhang Y. Suicide rates in China, 1995–99. *Lancet* 2002;**359**:835–40.
- 17 **Rao C**, Yang G, Hu J, *et al*. Validation of cause-of-death statistics in urban China. *Int J Epidemiol* 2007; doi: 10.1093/ije/dym003.
- 18 **Banister J**, Hill K. Mortality in China 1964–2000. *Popul Stud (Camb)* 2004;**58**:55–75.

- 19 Murray CJL, Lopez AD. *The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*. Cambridge, MA: Harvard School of Public Health on behalf of the World Health Organization and the World Bank, 1996.
- 20 Ruzicka L, Lopez AD. The use of cause-of-death statistics for health situation assessment: national and international experiences. *World Health Stat Q* 1990;**43**:249–57.
- 21 Kumar V, Datta N. Lay reporting and verbal autopsy in assessment of infant mortality. *Indian J Pediatr* 1986;**53**:672–4.
- 22 Datta N, Mand M, Kumar V. Validation of causes of infant death in the community by verbal autopsy. *Indian J Pediatr* 1988;**55**:599–604.
- 23 Mirza NM, Macharia WM, Wafula EM, et al. Verbal autopsy: a tool for determining cause of death in a community. *East Afr Med J* 1990;**67**:693–8.
- 24 Risse G. Cause of death as a historical problem. *Continuity Change* 1997;**12**:175–88.
- 25 Cook IG, Dummer TJ. Changing health in China: re-evaluating the epidemiological transition model. *Health Policy* 2004;**67**:329–43.
- 26 Prescott N, Jamison DT. The distribution and impact of health resource availability in China. *Int J Health Plann Manage* 1985;**1**:45–56.
- 27 Shaokang Z, Zhenwei S, Blas E. Economic transition and maternal health care for internal migrants in Shanghai, China. *Health Policy Plan* 2002;**17**(Suppl):47–55.
- 28 Surault P. A research into the determinants of socio-economic differences in health. *Rev Prat* 2004;**54**:2233–40.
- 29 Liang J, McCarthy JF, Jain A, et al. Socioeconomic gradient in old age mortality in Wuhan, China. *J Gerontol B Psychol Sci Soc Sci* 2000;**55**:S222–33.

Stay a step ahead with *Online First*

We publish all our original articles online before they appear in a print issue. This means that the latest clinical research papers go straight from acceptance to your browser, keeping you at the cutting edge of medicine. We update the site weekly so that it remains as topical as possible. Follow the Online First link on the home page and read the latest research.