

Selecting a measure of health-related quality of life.

Dempster, M., & Donnelly, M. (2000). Selecting a measure of health-related quality of life. *Social Work in Health Care*, 32(1), 45-56.

Published in:
Social Work in Health Care

Document Version:
Early version, also known as pre-print

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Selecting A Measure Of Health Related Quality Of Life

Martin Dempster & Michael Donnelly

Abstract

Quality of life is becoming recognized increasingly as an important outcome measure which needs to be considered by social workers. However, there does not appear to be a clear consensus about the definition of quality of life. In addition, social workers are likely to experience difficulties choosing and applying an appropriate instrument with which to measure quality of life because of the many available instruments purporting to assess quality of life. This paper discusses the definition of health-related quality of life and explains the main measurement properties of an instrument that must be appraised when considering whether or not an instrument is appropriate. The paper will assist social workers to make an informed choice about measures of health-related quality of life.

Key words: quality of life, measuring outcomes

This manuscript was authored by Dr Martin Dempster BSc PhD and Dr Michael Donnelly BSc PhD from the Health & Social Care Research Unit, Queen's University Belfast, Mulhouse Building, Grosvenor Road, Belfast, Northern Ireland, BT12 6BJ.

Tel: +44 1232 331463 Fax: +44 1232 320664

e-mail: m.dempster@qub.ac.uk

Please send all correspondence to Martin Dempster at the above address.

Health-related quality of life

Selecting A Measure Of Health Related Quality Of Life

Abstract

Quality of life is becoming recognized increasingly as an important outcome measure which needs to be considered by social workers. However, there does not appear to be a clear consensus about the definition of quality of life. In addition, social workers are likely to experience difficulties choosing and applying an appropriate instrument with which to measure quality of life because of the many available instruments purporting to assess quality of life. This paper discusses the definition of health-related quality of life and explains the main measurement properties of an instrument that must be appraised when considering whether or not an instrument is appropriate. The paper will assist social workers to make an informed choice about measures of health-related quality of life.

Social workers (and other care professions) are being encouraged increasingly to consider a clients' quality of life during the assessment process and to evaluate their practice or services using quality of life outcome-indicators, especially in community settings and services (Rosenberg & Holden, 1997). It is necessary for social workers to become aware of issues surrounding the measurement of quality of life in order to achieve these goals. This paper provides a description and discussion of the important issues that need to be considered when determining the most appropriate and dependable quality of life measure.

What Does "Quality of Life" Mean?

In basic terms, quality of life can be perceived as the degree to which human needs such as those outlined by Maslow (1943) are satisfied. These tend to be grouped into physical, spiritual, social, economic and psychological needs and, thus, quality of life is a label that covers a broad range of life "domains" or areas. However, it is more appropriate and practicable in the context of health and social care services to investigate a more specific definition of quality of life than the construct which encompasses the above noted broad definition. One specific focus is health-related quality of life (HRQL), which refers to the measurement of the quality of those aspects of a person's life that impact directly upon their health (Patrick & Erickson, 1993).

Most attempts to define HRQL tend to be based on the statement by the World Health Organization (WHO) (Guyatt, 1993). The WHO (1958) declared that "health

Health-related quality of life

is a state of complete physical, mental and social wellbeing and not merely the absence of disease or infirmity". This statement introduced the concept of "positive health" and the notion that an assessment of health should include psychological and social factors in addition to the traditional outcomes of mortality and morbidity. There appears to be general agreement that this holistic approach should be applied to health assessment and that any assessment of HRQL should consider psychological, social and physical factors (Martin & Stockler, 1998).

However, a consensus is lacking regarding the nature, content and measurement of the specific psychological, social and physical factors that constitute HRQL and how these factors should be combined to provide an index of HRQL (Avis & Smith, 1994). For example, two of the instruments most commonly used to measure HRQL are the Short Form 36 (SF-36), developed by Ware *et al.* (1993) and the Nottingham Health Profile Part I (NHP), authored by Hunt *et al.* (1980). The SF-36 assesses HRQL by measuring eight factors or domains: physical functioning, social functioning, vitality, bodily pain, mental health, general health, role limitations due to physical problems and role limitations due to emotional problems. The NHP Part I assesses HRQL by measuring 6 domains: mobility, pain, energy, sleep, emotional reactions and social isolation. A comparison of the content of these two instruments reveals that the authors of the NHP considered sleep to be important in an assessment of HRQL. The SF-36 does not include this domain and, furthermore, incorporates other domains which are not covered by the NHP Part I such as the performance of household tasks, work and family relationships. Therefore, the assessment of HRQL may vary according to the instrument chosen to measure HRQL. Researchers should state clearly the definition that is being used to measure HRQL as

well as the nature and method of assessment. Gill & Feinstein (1994) suggested that researchers should also justify their reasons for choosing a particular instrument.

This practice is similar to one of the stages of scientific research involving the formulation of an operational definition. Nevertheless, there are few reports in the research literature which describe attempts to operationally define quality of life (Bowling, 1995). Gill & Feinstein (1994) discovered, in a random sample of 75 HRQL articles, that less than 47% of articles identified the HRQL domains of interest and 52% offered no explanation for their choice of instrument. The practice of justifying the selection of a HRQL instrument within a study is likely to facilitate the translation of research into practice and assist users of research to draw appropriate conclusions. This facilitative process will be enhanced further when social workers develop an awareness of the properties of an instrument that need to be appraised in order to select a measure of HRQL.

Why Measure Health-Related Quality of Life?

The measurement of HRQL, in addition to physiological measures, provides a more complete assessment of a person's health and well-being (Wood-Dauphinee, 1999). It may be argued that social workers "measure" HRQL in the course of their work when they consider the effects of an illness or its treatment on the psychological, social and physical aspects of a person's life even though they may not use formal HRQL instruments, employ the label HRQL, or create a numerical index to grade HRQL. However, there are advantages to using a formal approach to HRQL

Health-related quality of life

measurement. One advantage concerns the need for scientific rigor when measuring health related outcomes, especially when treatment decisions can be affected by these outcomes (Muldoon *et al.*, 1998). It is essential therefore that instruments are reliable, valid and, overall, psychometrically sound. Secondly, it is important to obtain a measure of HRQL from the perspective of the client (Smith & Avis, 1997). It is difficult for social workers to determine how an illness or an intervention will impinge on the life of each individual; and clinicians' ratings of a client's quality of life are often very different from the ratings provided by clients (Jachuck *et al.*, 1982). This has been recognized in the field of medicine: “What matters...is how the patient feels, rather than how doctors think they ought to feel on the basis of clinical measurements” (Bowling, 1991). Finally, an investigation of HRQL by a social worker may enhance client-social worker communication and provide clients with a welcome opportunity to reveal problems which otherwise might be overlooked (Martin & Stockler, 1998). As Berkman *et al.* (1999) state: “There is a growing recognition of the need to go beyond traditional physiologic variables to include the other psychosocial components of HRQL that, of course, are of importance to social work practice.”

Indeed, it appears that we have now reached the stage whereby anyone undertaking an evaluation of treatment or practice must justify any decision to exclude a measure of HRQL (Bausell, 1998). The next section outlines the measurement properties of instruments which social workers should scrutinize in the course of making a decision about the most appropriate assessment strategy.

How to Measure Health-Related Quality of Life

Unfortunately the degree of interest in HRQL has not been matched by a commensurate level of expertise in its measurement (Barnett, 1991). There is a need to use measures of HRQL that have been subjected to rigorous testing and examination. If it is important to measure HRQL then efforts must be directed towards ensuring that it is measured properly.

Generic and condition-specific measures

A point to consider when measuring HRQL is whether to use a generic or a condition-specific instrument. A generic instrument is one that attempts to measure a broad range of domains that are related to HRQL. As generic instruments cover a variety of areas, they are used to index HRQL within different client populations. This is perceived as an advantage as it allows comparisons of outcomes to be made between client groups. Condition-specific instruments are designed to measure HRQL by tapping those areas of life that are particularly pertinent for clients with a specific condition. Although these instruments have a narrow focus, they have a couple of advantages over their generic counterparts. Firstly, condition-specific instruments will explore areas which are related closely to the areas of life explored by social workers and secondly, condition-specific instruments are likely to be more sensitive to change in health status than generic instruments (Tullis & Guyatt, 1995). Often it is recommended that both types of instruments should be used when evaluating HRQL.

Is the HRQL instrument to be used for descriptive or evaluative purposes?

HRQL measurements can be used both to describe the HRQL of clients with a condition (and so discriminate between groups of clients) and to evaluate the impact of a service or treatment. A HRQL instrument needs to be administered only to the study participants at one point in time when the purpose is to describe or differentiate between clients. For example, social workers might want to explore the differences in HRQL between men and women with human immunodeficiency virus who have access to social work services (Davidson *et al.*, 1998). Alternatively, social workers might want to assess the impact on HRQL of a social work and medical intervention for older people with eye disorders (Isralowitz *et al.*, 1996). This study has an evaluative purpose and a repeated measures design is required where the HRQL instrument is administered at more than one point in time – usually before and after the introduction of the intervention.

The selection of an appropriate instrument will be determined by whether the purpose of a study is descriptive (discriminative) or evaluative (Fletcher & Bulpitt, 1988). Discriminative and evaluative instruments need to be reliable and valid; evaluative instruments must also be sensitive (Guyatt *et al.*, 1992). Assessments of reliability, validity and sensitivity are the basis of any psychometric investigation but others have proposed that consideration should also be given to interpretability, respondent and administrative burden, and cultural and language adaptations (Nanda & Andresen, 1998).

Assessing validity

In basic terms, validity is the ability of an instrument to measure accurately what it claims to measure. In practice there are several types of validity. Face validity is the most basic and is a subjective assessment of whether or not the instrument appears to be able to measure what it claims to measure. Criterion-related validity assesses the relationship between the instrument under investigation and a criterion which is usually a “gold standard” – a measurement instrument which has been tried and tested and is accepted as giving an accurate measurement of the variable of interest. However, as the definition of HRQL is dynamic there is no definitive “gold standard” and so there is a need to choose a criterion measure which has been shown to be psychometrically sound and which measures similar domains of HRQL as the instrument under scrutiny. There are two types of criterion-related validity:

1. Concurrent validity where the instrument under investigation and the criterion are administered at the same point in time.
2. Predictive validity where the data from the administration of the instrument under scrutiny is correlated with the data from a criterion measure available at a future point in time.

In both cases, the data from the instrument under investigation and the “gold standard” instrument are correlated to provide a validity coefficient. A strong correlation coefficient (usually 0.7 or greater) indicates strong concurrent validity. There is little consensus about the value of the correlation coefficient that constitutes a cut-off point for strong validity, so an instrument cannot be said to be valid or invalid. Rather, validity should be perceived as a continuous measure and the stronger the correlation coefficient, the stronger is the evidence for the validity of the

instrument. As an example, consider the results for two HRQL instruments used among people with ischemic heart disease. The Seattle Angina Questionnaire (Spertus *et al.*, 1995) and the Quality of Life Index (Ferrans & Powers, 1985) both contain a domain which measures physical functioning. Dougherty *et al.* (1998) assessed the validity of these domains by correlating them with the SF-36 physical functioning scale (the gold standard). They found that the Seattle Angina Questionnaire physical functioning domain and the SF-36 physical functioning domain had a correlation of 0.63; the Quality of Life Index physical functioning domain and the SF-36 physical functioning domain had a correlation of 0.52. Therefore, if Dougherty *et al.* were correct to treat the SF-36 scale as a gold standard, it could be concluded that there was stronger evidence for the concurrent validity of the physical functioning domain of the Seattle Angina Questionnaire, compared to the Quality of Life Index.

Finally, the investigation of construct validity is becoming increasingly popular as a result of the development of user-friendly computer packages that facilitate the use of statistical techniques such as factor analysis. Construct validity investigates the extent to which the data collected from an instrument conform to the hypothetical constructs that underlie the measurement instrument. It is an ongoing process of testing that can influence the development of the underlying constructs and theories as well as the development of measurement instruments. In a clinical setting one of the most important aspects of construct validity is discriminative validity. In other words, the ability of an instrument to distinguish between clinically different groups. Again, this can be demonstrated with reference to people with heart disease. McCarthy *et al.* (1995) reports that scores on the SF-36 scales were significantly

lower for people awaiting cardiac surgery than for the general population, thereby providing evidence for the discriminative validity of the SF-36.

Assessing reliability

Reliability refers to the consistency of an instrument and this can have two meanings: consistency over repeated testing or internal consistency. Consistency over repeated testing is commonly referred to as test-retest reliability. This involves administering the instrument at one point in time and then administering the instrument to the same group of respondents, whose condition has remained stable, at a second point in time. The scores from each administration are correlated and the resulting correlation coefficient is an estimate of the instrument's test-retest reliability. If respondents attain exactly the same results at both points in time, the test-retest reliability coefficient will be 1.00. In practice this is highly unlikely but the closer the reliability coefficient is to 1.00, the more reliable is the instrument. For example, Carver *et al.* (1999) report a test-retest reliability coefficient of 0.96 for the SF-20 scale (Stewart *et al.*, 1988) when administered to older people living at home. This would be considered a very high reliability coefficient. The problem with using this approach to assess HRQL measurement instruments is that the health of some individuals may significantly improve or deteriorate over time and so the time delay between the two administrations of the instrument must be kept short.

An instrument is said to have high internal consistency if the items which are supposed to tap into a particular domain all correlate highly with one another. This is based on the notion that if all of these items are supposed to be measuring the same

concept, then responses on these items should all be of similar magnitude. Estimates of internal consistency are commonly made using the Cronbach's Alpha statistic. Andresen *et al.* (1998) recommends the use of the SF-36 rather than the Sickness Impact Profile (Bergner *et al.*, 1976) to measure HRQL among community-living older adults and cites as supporting evidence the Cronbach's Alpha statistic for the two scales – the Sickness Impact Profile ranged from 0.59 to 0.87; the SF-36 ranged from 0.69 to 0.93. The minimum acceptable level for the demonstration of high internal consistency is often accepted to be 0.80 (Bryman & Cramer, 1997).

Assessing responsiveness

Information regarding the sensitivity or responsiveness to change of an instrument is not reported as often as information about reliability and validity. This is surprising given the importance of this characteristic when instruments are to be used for evaluative purposes (Guyatt *et al.*, 1987) but may be explained by the fact that there are several ways of estimating sensitivity and it could be the lack of consensus about the definition and determination of this concept which leads to a relative lack of information on assessment methods, compared to the vast amount of information regarding the assessment of reliability and validity.

The two most commonly cited statistics used in the assessment of sensitivity are the effect size and the standardized response mean. The effect size is the ratio of change in average scores over time to the standard deviation of scores at baseline $((\text{mean at time 1} - \text{mean at time 2}) / \text{standard deviation at time 1})$ (Fitzpatrick *et al.*, 1992). This equation allows us to estimate sensitivity or responsiveness by calculating

the magnitude of change in relation to the between-subject variability. The standardized response mean uses the standard deviation of the change scores as the denominator (Garratt *et al.*, 1994). Some authors have advocated the standardized response mean because it focuses attention upon the distribution of change scores, which should be the focal point when assessing responsiveness to change (Harper *et al.*, 1997), however it is problematic to calculate when there are unequal numbers at the two points in time. Therefore, the standardized response mean should be used when the same people complete the instrument at both points in time and the effect size should be used when there is data for respondents at only one of the two points in time. Standardized response means and effect sizes are interpreted in the same manner – small moderate and large changes are represented by the values 0.2, 0.5 and 0.8, respectively (Liang *et al.*, 1990).

The Geriatric Quality of Life Questionnaire was developed by Guyatt *et al.* (1993). One of its scales measures activities of daily living. Guyatt *et al.* report that this scale has a standardized response mean of 0.26 and although this is low, it is higher than the reported standardized response mean for the Barthel Index (0.20).

Conclusion

In summary, when choosing a measure of HRQL it is important to ensure that it addresses the domains that are relevant to particular client groups and that it is supported by good psychometric evidence. It is important to emphasize that when investigating validity, reliability and sensitivity, one should not expect a dichotomous answer. In other words, it is almost impossible to answer the question - is this

instrument valid – with a yes or no response. Indeed, the psychometric assessment of any instrument is an ongoing process. This is why a comparative analysis of instruments is a useful process – it allows one to determine which instrument has the strongest psychometric evidence.

The recognition of HRQL as an important outcome indicator has resulted in HRQL evaluation receiving an increasing level of attention by researchers, practitioners and clinicians. It is likely that the sheer abundance of instruments that are available means that the measurement of HRQL is difficult to access and use by many social workers. This paper has attempted to provide a set of guidelines which social workers may use to scrutinize HRQL instruments and so make an appropriate choice of measure for their purposes. It is essential that the psychometric results of HRQL instruments that are used among different client groups in social work practice are disseminated in order to base the choice of instrument on best available evidence.

References

- Andresen, E.M., Rothenberg, B.M., Panzer, R., Katz, P., & McDermott, M.P. (1998). Selecting a generic measure of health-related quality of life for use among older adults. Evaluation & the Health Professions, 21(2), 244-264.
- Avis, N.E., & Smith, K.W. (1994). Conceptual and methodological issues in selecting and developing quality of life measures. In G. L. Albrecht & R. Fitzpatrick (Eds.), Advances in Medical Sociology. (pp. 255-280). Greenwich: JAI Press.
- Barnett, D.B. (1991). Assessment of quality of life. American Journal of Cardiology, 67(12), 41C-44C.
- Bausell, R.B. (1998). Quality of life assessment in outcomes research. Evaluation & the Health Professions, 21, 139-140.
- Bergner, M., Bobbitt, R.A., Pollard, W.E., & et al. (1976). The Sickness Impact Profile: validation of a health status measure. Medical Care, 14, 57-67.

Berkman, B., Chauncey, S., Holmes, W., Daniels, A., Bonander, E., Sampson, S., & Robinson, M. (1999). Standardized screening of elderly patients' needs for social work assessment in primary care: use of the SF-36. Health & Social Work, 24, 9-16.

Bowling, A. (1991). Measuring Health. A Review of Quality of Life Measurement Scales. Milton Keynes: Open University Press.

Bowling, A. (1995). What things are important in people's lives? A survey of the public's judgements to inform scales of health related quality of life. Social Science and Medicine, 41, 1447-1462.

Bryman, A., & Cramer, D. (1997). Quantitative Data Analysis. London: Routledge.

Carver, D.J., Chapman, C.A., Thomas, V.S., Stadnyk, K.J., & Rockwood, K. (1999). Validity and reliability of the Medical Outcomes Study Short Form-20 questionnaire as a measure of quality of life in elderly people living at home. Age and Ageing, 28, 169-174.

Davidson, A.J., Bertram, S.L., Lezotte, D.C., Marine, W.M., Rietmeijer, C.A., Hagglund, B.B., & Cohn, D.L. (1998). Comparison of health status, socioeconomic characteristics, and knowledge and use of HIV-related resources between HIV-infected women and men. Medical Care, 36, 1676-1684.

Dougherty, C.M., Dewhurst, T., Nichol, W.P., & Spertus, J. (1998). Comparison of three quality of life instruments in stable angina pectoris: Seattle Angina Questionnaire, Short Form Health Survey (SF-36), and Quality of Life Index-Cardiac Version III. Journal of Clinical Epidemiology, 51(7), 569-575.

Ferrans, C.E., & Powers, M.J. (1985). Quality of Life Index: development and psychometric properties. Advances in Nursing Science, 8(1), 15-24.

Fitzpatrick, R., Ziebland, S., Jenkinson, C., & Mowat, A. (1992). Importance of sensitivity to change as a criterion for selecting health status measures. Quality in Health Care, 1, 89-93.

Fletcher, A.E., & Bulpitt, C.J. (1988). Measurement of quality of life in clinical trials of therapy. Cardiology, 75 Suppl 1, 41-52.

- Garratt, A.M., Ruta, D.A., Abdalla, M.I., & Russell, I.T. (1994). SF 36 health survey questionnaire: II. Responsiveness to changes in health status in four common clinical conditions. Quality in Health Care, 3, 186-192.
- Gill, T.M., & Feinstein, A.R. (1994). A critical appraisal of the quality of quality-of-life measurements. Journal of the American Medical Association, 272(8), 619-626.
- Guyatt, G.H., Eagle, D.J., Sackett, B., Willan, A., Griffith, L., McIlroy, W., Patterson, C.J., & Turpie, I. (1993). Measuring quality of life in the frail elderly. Journal of Clinical Epidemiology, 46(12), 1433-1444.
- Guyatt, G.H., Kirshner, B., & Jaeschke, R. (1992). Measuring health status: what are the necessary measurement properties? Journal of Clinical Epidemiology, 45, 1341-1345.
- Guyatt, G.H., Walter, S., & Norman, G. (1987). Measuring change over time: assessing the usefulness of evaluative instruments. Journal of Chronic Diseases, 40, 171-178.

- Harper, R., Brazier, J.E., Waterhouse, J.C., Walters, S.J., Jones, N.M.B., & Howard, P. (1997). Comparison of outcome measures for patients with chronic obstructive pulmonary disease (COPD) in an outpatient setting. Thorax, *52*, 879-887.
- Hunt, S.M., McKenna, S.P., McEwen, J., Backett, E.M., Williams, J., & Papp, E. (1980). A quantitative approach to perceived health status: a validation study. Journal of Epidemiology and Community Health, *34*, 281-286.
- Isralowitz, R.E., Rosenthal, G., Lifshitz, T., Madar, M., & Shabtay, M. (1996). Ethiopian elderly immigrants: a community-based social work and medical intervention to address eye care needs. Journal of Gerontological Social Work, *27*, 167-177.
- Jachuck, S.J., Brierley, H., Jachuck, S., & Willcox, P.M. (1982). The effect of hypotensive drugs on the quality of life. Journal of the Royal College of General Practitioners, *32*, 103-105.
- Liang, M.H., Fossel, A.H., & Larson, M.G. (1990). Comparisons of five health status instruments for orthopedic evaluation. Medical Care, *28*, 632-642.

Martin, A.J., & Stockler, M. (1998). Quality of life assessment in health care research and practice. Evaluation & the Health Professions, 21(2), 141-156.

Maslow, A.H. (1943). A theory of human motivation. Psychological Review, 50, 370-396.

McCarthy, M.J.J., Shroyer, A.L., Sethi, G.K., Moritz, T.E., Henderson, W.G., Grover, F.L., London, M.J., Gibbs, J.O., Lansky, D., & Miller, D. (1995). Self-report measures for assessing treatment outcomes in cardiac surgery patients. Medical Care, 33(10 Suppl), OS76-OS85

Muldoon, M.F., Barger, S.D., Flory, J.D., & Manuck, S.B. (1998). What are quality of life measurements measuring? British Medical Journal, 316, 542-545.

Nanda, U., & Andresen, E.M. (1998). Health-related quality of life - A guide for the health professional. Evaluation & the Health Professions, 21, 179-215.

Patrick, D.L., & Erickson, P. (1993). Assessing health-related quality of life for clinical decision making. In S. R. Walker & R. M. Rosser (Eds.), Quality of Life Assessment. Key Issues in the 1990s . (pp. 11-64). Lancaster: Kluwer.

Smith, K.W., & Avis, N.E. (1997). Measuring quality of life. Medical Principles and Practice, 6, 171-176.

Spertus, J.A., Winder, J.A., Dewhurst, T.A., Deyo, R.A., Prodzinski, J., McDonnell, M., & Fihn, S.D. (1995). Development and evaluation of the Seattle Angina Questionnaire - a new functional status measure for coronary artery disease. Journal of the American College of Cardiology, 25, 333-341.

Stewart, A.L., Hays, R.D., & Ware, J.E. (1988). The MOS Short Form general health survey: reliability and validity in a patient population. Medical Care, 26, 724-735.

Tullis, D.E., & Guyatt, G.H. (1995). Quality of life in Cystic Fibrosis. Pharmacoeconomics, 8, 23-33.

Ware, J.E., Snow, K.K., Kosinski, M.K., & Gandek, B. (1993). SF-36 Health Survey Manual and Interpretation Guide. Boston: The Health Institute, New England Medical Center.

Wood-Dauphinee, S. (1999). Assessing quality of life in clinical research: from where have we come and where are we going? Journal of Clinical Epidemiology, 52, 355-363.

World Health Organization (1958). The First Ten Years of the World Health Organization. Geneva: World Health Organization.