Appendix D Statistical methods

Comparisons and tests of statistical significance

This report includes statistical tests of the significance of comparisons of rates between population groups. Any statistical comparison applied to one variable must take account of any other potentially relevant variables. For example, any comparison of participation by state must also take account of differences in the distribution of age and sex between the states. These other variables are known as confounding variables.

Crude rates

A crude rate is defined as the number of events over a specified period divided by the total population. The crude rate (for participation, attendance and follow-up) is the proportion of people who have proceeded to a key point on the screening pathway (at the date of the data extraction) out of those eligible to proceed to that point. For example, the crude participation rate is the proportion of the eligible people invited in 2008 who return a completed FOBT kit by 31 January 2009. The crude colonoscopy follow-up is the proportion of people invited in 2008 with a positive FOBT result who proceeded to colonoscopy by 31 January 2009.

The crude proportions will generally underestimate the true proportions of the population who participated in the NBCSP. This is because at any point in time there are members of the population who are eligible to proceed to the next point on the screening pathway, but who have not yet had time to do so. For example, a person who has just received an invitation to screen may intend to participate in screening but may not have had time to do so. They will be counted in the denominator of the crude participation but not in the numerator. Similarly, there is a time lag between when a person with a positive FOBT result is referred for colonoscopy and when they can actually have the colonoscopy. A colonoscopy follow-up calculated during this lag includes them in the denominator but not in the numerator.

Age-specific rates

Age-specific rates were calculated by dividing the number of cases occurring in each specified age group by the corresponding population in the same age group, expressed as per 100,000 persons.

Age-standardised rates

Rates are adjusted for age to help comparisons between populations that have different age structures; for example, between youthful and ageing communities. Two different methods are commonly used to adjust for age. In this publication direct standardisation was used, in which age-specific rates were multiplied against a constant population (the Australian 2001 population). This effectively removes the influence of age structure on the summary rate, and is described as the age-standardised rate. The method used for this calculation comprises three steps:

- Calculate the age-specific rate for each age group.
- Calculate the expected number of cases in each 5-year age group by multiplying the age-specific rates by the corresponding standard population, and dividing by 100,000, giving the expected number of cases.
- Calculate the age-standardised rate by summing the expected number of cases in each age group, and dividing this sum by the total of the standard population used in the calculation and multiplying by 100,000.

Confidence intervals

Different methods of calculating 95% confidence intervals were used for the crude rates in NBCSP chapters 1–5 (Box D.1), and the age-standardised rates presented in the incidence and mortality chapters (Box D.2). These confidence intervals indicate the variation that might be expected in such estimates purely by chance.

Box D.1: Confidence intervals for proportions

Confidence intervals for crude proportions (p) in chapters 1–5 were calculated using the basic confidence interval formula for binomial proportions:

$$95\% \text{ CI for proportions} = p \pm 1.96 \times \sqrt{\frac{p \times (1-p)}{\text{Number of cases}}}$$

Box D.2: Confidence intervals for age-standardised rates

The confidence intervals for age-standardised rates in the incidence and mortality chapters were calculated using the methods presented by Holman et al. (1987). A relatively simple approximation of the confidence intervals that readers might use when looking at age-standardised rates is:

$$95\% \text{ CI approximation} = \text{AS rate} \pm 1.96 \times \frac{\text{ASR}}{\sqrt{\text{Number of cases}}}$$

Kaplan-Meier estimates of participation and follow-up

The Bowel Cancer Screening Pilot Program employed the use of Kaplan-Meier estimates of participation, attendance and follow-up. This statistical method calculates a modelled rate based on the time it takes each individual invited for screening to move between points on the screening pathway. For example, participation is calculated by following each invited person and, for those who respond, recording the time it takes them to respond. This allows the calculation of a response rate over time from the date of invitation. Kaplan-Meier methods are standard methods used to model the time to an event and the changes in the rates of an event over time. In this case, the event is a person's response (by returning a completed FOBT kit), and the time to the event is measured in weeks from the date the invitation was sent. These Kaplan-Meier estimates represent valid estimates of the true FOBT participation. The use of Kaplan-Meier estimates in the NBCSP was endorsed by the

Implementation Advisory Group, and allows direct comparison of participation, attendance and follow-up rates with the Bowel Cancer Screening Pilot Program.

In principle, the Kaplan-Meier estimate only gives a result at a specific point in time. The estimate is likely to grow for later points in time. However, inspection of these estimates shows that they reach a plateau, after which they have only a negligible increase. Kaplan-Meier estimates in this report were calculated for participation at 52 weeks and primary health care practitioner and colonoscopy follow-up at 26 weeks. Further, preliminary analyses based on modelling the survival time with both a Weibull and an exponential distribution showed that the latest observed Kaplan-Meier estimate differed from the long-term modelled estimate by less than 1 percentage point. Hence, the latest Kaplan-Meier estimate can be taken as an approximate estimate of the overall rate.

The Kaplan-Meier estimates require that classifying variables be known for the population. Hence, they can be calculated for participation classified by age, sex and state. However, they cannot be used for participation classified by Aboriginal and Torres Strait Islander status, language group, or disability status, which are not known for all the invited population. These variables are only known for those participants who identify themselves as a member of these groups on their returned Participant Details form. In these cases, a crude participation can be calculated by using known population counts (from the Australian Bureau of Statistics Census data) in the denominator. However, the Kaplan-Meier estimates cannot be applied in this situation. In these cases, all analyses will be based solely on the crude participation. Therefore, the participation presented in this report for Aboriginal and Torres Strait Islander peoples, people with a disability and people with a language other than English may represent underestimates of the true proportions.

Aboriginal and Torres Strait Islander status, language group status and disability status will be known for all people completing FOBT kits (at least to the extent that people self-identify as members of these groups). Hence, in principle, Kaplan-Meier estimates can be calculated for these groups for participation at subsequent points on the screening pathway. In practice, these calculations depend on sufficient numbers of people self-identifying as group members to allow the calculation of reliable estimates.

Positive predictive value

The calculation of an accurate positive predictive value for the screening test relies on completeness of both colonoscopy and histopathology data. Due to the time lags between the receipt of a positive FOBT and colonoscopy and histopathology procedures, and under-reporting by clinicians, positive predictive values have not been calculated.