Detailed model description

### Model parameters

The survival rates post-surgery were obtained from national registries for oncological1 and cardiothoracic2 surgeries. For the remaining surgeries, data was obtained from scientific literature. The survival data pre-surgery for all surgeries is based on data from published studies. If either survival with or without treatment was lacking, the reported treatment effect (preferably from a randomized controlled trial) was used to calculate the missing survival parameter. The disease specific mortality was added to the overall age-specific mortality from the Central Bureau of Statistics in the Netherlands.3 The mean age of the patients was obtained from published studies. All survival data had to be converted to mortality risk per week (formulas presented in Appendix E).4

The QoL before and after surgery were based on ‘disutility weights’ from the Global Burden of Disease Study 2016.5 This study reports disability weights for nonfatal health conditions. These weights represent the magnitude of health loss associated with the conditions, where 0 represents no loss (full health) and 1 all lost (death). When these weights are multiplied with the duration lived in this conditions, one has calculated the weighted ‘years lived with disability’ (YLD).6 The YLD summed with the years of life lost to premature death (YLLs) give the disability adjusted life years (DALY).7 A ‘full DALY’ can be thought of as losing one year in full health. Disability Adjusted Life Years (DALYS) are the complement (the opposite) of the Quality Adjusted Life Years (QALYs), which represents the value of a year spent in full health. For our study, the complement (1-x) of the disability weight was used as QoL values to calculate QALYs.

Where possible, we based the QoL of health conditions directly on the GBD study data. The remaining conditions were estimated using methods described by Stouthard et al.8 We used a visual analogue scale (VAS) calibrated with GBD 2016 QoL weights. Stouthard et al. describe how experts can then place (map) the remaining health conditions on the VAS scale with QoL weights. Our protocol was slightly different form the protocol of Stouthard, in the way that we did not make use of the EQ-5D to classify all health conditions at hand. The expert panel consisted of a diverse group of healthcare professionals, both surgeons (e.g. cardiothoracic surgeons, neurosurgeons, and gynecologic surgeons) as well as generalists (e.g. internists, geriatricians and GPs). We ensured that there was a representative from the specialty of each surgery. The health conditions were valued one by one using the following procedure. First, the health condition was shortly introduced by an expert with the most clinical experience with this condition. The other experts were allowed to ask questions and discuss the QoL aspects of the condition. Subsequently, all experts wrote down their own QoL estimation of the health condition. Then, two to three other experts were then invited to express their estimated QoL value for the health condition. Ultimately, the expert registered their own final values. In this way, the expert could use a maximum of information and opinions, but still express their own estimation. In addition, we could estimate the variance, 95% confidence interval (95% CI), of the QoL values. The mean and 95% CI of the mapped QoL scores were used in the model. We used two session of three hours to collect all QoL value. The preoperative and postoperative health state of 3 surgeries (one with a mild and severe subgroup) were estimated in both sessions, which effectively were 8 estimates of QoL. This allows us to get an indication of the reliability (test-retest by means of a t-test) of the valuations. For the model, the first estimates obtained in the first session were used. Appendix D provides the calibrated VAS as well as an overview of the expert panel.

Since postponing surgery can have consequences on the effectiveness of the surgery, we included a model parameter that reflected the time until no effect can be expected of treatment on survival. In practice, this means that when this time has passed, we assumed that the surgery did no longer have an effect on the survival of the patient anymore. This time is often important in oncological surgeries, where after a specific time a tumor becomes inoperable or metastasize. The effectiveness of types of surgery, for example repairing an abdominal aneurysm of the aorta, could be time-dependent as well. The data for this parameter was obtained from the scientific literature (Appendix A). For most surgeries, only data about the minimal delay not associated with worse survival could be obtained from the scientific literature. For those surgeries, we assumed the upper limit of this parameter to be a year (the maximum delay we evaluated), and the mean of the lower and upper limit as average. The same was done for the time until no effect can be expected on QoL.

### Assumptions

The design of the model translates to the following core assumptions:

* The health benefit of the surgery for the average patient is evaluated, which means that the model does not take into account individual patient characteristics, prognostic factors or co-morbidities.
* The model does not include complications or a period of recovery, both of which can reduce QoL temporarily.
* Surgeries were assumed to be successful without complications.
* No increased risk of mortality during surgery associated with delay is assumed.
* The COVID-19 context does not impact the performance of the surgeries.
* Complications and harm associated with surgery do not differ between various delays. Therefore, the measures of urgency, QALY and LY loss per month, can be compared across treatments with varying associated harm.

Since benefits now are enjoyed more that in the distant future, it is recommended to perform discounting.9,10 A discount rate of 0.015 per year for health benefits was used, as this is common practice in the Netherlands.11 Discounting makes current benefits worth more than those expected in the future. If discounting is not performed, we would value health gains achieved this year equal to those achieved in 30 or 40 years.

1. Kankersoorten - IKNL [Internet]. [cited 2020 May 19];Available from: https://iknl.nl/kankersoorten

2. NHR [Internet]. [cited 2020 May 19];Available from: https://nederlandsehartregistratie.nl/

3. CBS. Sterftekansen naar leeftijd, geslacht, opleidingsniveau [Internet]. [cited 2020 May 19];Available from: https://www.cbs.nl/nl-nl/maatwerk/2017/23/sterftekansen-naar-leeftijd-geslacht-opleidingsniveau

4. Hunink M, Mc E, Glasziou P, Elstein A. Decision Making in Health and Medicine: Integrating Evidence and Values [Internet]. 2nd ed. Cambridge: Cambridge University Press; 2003 [cited 2020 May 19]. Available from: http://www.cambridge.org

5. Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2016 (GBD 2016) Disability Weights. Seattle, United States: Institute for Health Metrics and Evaluation (IHME): 2017.

6. General Guidance for DALYs calculation [Internet]. [cited 2020 May 19]. Available from: https://montagu.vaccineimpact.org/contribution/resources/c978e5a1acf6a502679c92200e78ef61.pdf

7. Disability-adjusted life years (DALYs) [Internet]. [cited 2020 May 19];Available from: https://www.who.int/data/gho/indicator-metadata-registry/imr-details/158

8. Stouthard EA, Essink-Bot M-L, Bonsel GJ. Disability weights for diseases A modified protocol and results for a Western European region. Eur J Public Health [Internet] 2000 [cited 2020 May 14];10(1):24–30. Available from: https://academic.oup.com/eurpub/article-abstract/10/1/24/490779

9. Torgerson DJ, Raftery J. Economic notes. Discounting. BMJ [Internet] 1999;319(7214):914–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10506056

10. Zorginstituut Nederland. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg. 2016;

11. Zorginstituut Nederland. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg. 2016.