Impact of surgical-site infection on health utility values: a meta-analysis

Agi M. McFarland1,2*, Sarkis Manoukian1,2, Helen Mason1,2 and Jacqui S. Reilly1,3

1Faculty of Health Sciences and Sport, University of Stirling, Stirling, UK
2Glasgow Caledonian University Yunus Centre for Social Business, Glasgow
3Health and Safeguarding Health through Infection Prevention (SHIP) Research Group, Glasgow

*Correspondence to: Agi M. McFarland, Faculty of Health Sciences and Sport, University of Stirling, Stirling FK9 4LA, UK (e-mail: agi.mcfarland@stir.ac.uk)

Abstract

Background: Surgical-site infections (SSIs) are recognized as negatively affecting patient quality of life. No meta-analysis of SSI utility values is available in the literature to inform estimates of this burden and investment decisions in prevention.

Methods: A systematic search of PubMed, MEDLINE, CINAHL, and the National Health Service Economic Evaluation Database was performed in April 2022 in accordance with PROSPERO registration CRD 42021262633. Studies were included where quality-of-life data were gathered from adults undergoing surgery, and such data were presented for those with and without an SSI at similar time points. Two researchers undertook data extraction and quality appraisal independently, with a third as arbiter. Utility values were converted to EuroQol 5D (EQ-5D™) estimates. Meta-analyses were conducted using a random-effects model across all relevant studies, with subgroup analyses on type and timing of the SSI.

Results: In total, 15 studies with 2817 patients met the inclusion criteria. Six studies across seven time points were used in the meta-analysis. The pooled mean difference in EQ-5D™ utility in all studies combined was –0.08 (95 per cent c.i. –0.11 to –0.05; prediction interval –0.16 to –0.01; I² = 40 per cent). The mean difference in EQ-5D™ utility associated with deep SSI was –0.10 (95 per cent c.i. –0.14 to –0.06; I² = 0 per cent) and the mean difference in EQ-5D™ utility persisted over time.

Conclusion: The present study provides the first synthesized estimate of SSI burden over the short and long term. EQ-5D™ utility estimates for a range of SSIs are essential for infection prevention planning and future economic modelling.

Introduction

Healthcare-associated infections (HAIs) are globally the most frequently reported adverse events. Surgical-site infection (SSI) is the most surveyed and frequent HAI in low- and middle-income countries, and the second most frequent HAI in Europe and the USA1. Once developed, SSIs present a significant burden in terms of disability and excess duration of hospital stay2, and contribute to the development of antimicrobial resistance through the increased use of antimicrobials3. Postoperative recovery is extended by SSI and increased use of healthcare resource is required. This impact on morbidity and mortality has a significant effect on patient quality of life4.

A recent systematic review5 highlighted the negative impact of SSI across six themes of patient functioning, namely physical, psychological, social, spiritual, economic, and the health care worker–patient relationship. Patients who develop SSI after surgery have a longer hospital stay by up to 24 days according to some estimates6, with an even greater impact in older people and those infected with resistant microbial strains. Deep SSIs have been reported to be associated with substantial levels of suffering and pain, which may become lifelong7. SSIs are also associated with poorer clinical pain and functional outcomes in a variety of surgical categories8,9. In an attempt to reduce the burden of such infections, SSI prevention is a mainstay of surgical patient care at both national10 and international11 levels.

Economic evaluations of interventions are used to inform investment decisions in healthcare within a constrained-resource context12. Cost–utility analysis, a specific form of cost-effectiveness analysis whereby effectiveness is measured through utility13, is ideally suited to explore the economic implications of SSI prevention strategies. Utility can be considered a health valuation on a scale anchored at 0 for dead and 1 for perfect health. This allows for a common outcome unit for interventions that affect both morbidity and mortality14, such as SSI. Utility values can subsequently be used to calculate quality-adjusted life years (QALYs), which are the standard health outcome measure sought by the National Institute for Health and Care Excellence for any interventions commissioned for the National Health Service (NHS) and other public sector bodies in the UK15. A recent systematic review15 identified a paucity of high-quality economic evaluations examining the benefits of SSI prevention. Synthesized utility values for SSI that could be used to generate such evidence are also lacking.

Therefore, a systematic review was undertaken to identify and summarize the utility decrement associated with SSI in adults undergoing surgery reported in the extant evidence base. This information can be used to both quantify the...
excess burden posed by SSI as well as for future cost–utility analyses evaluating the cost effectiveness of SSI prevention measures. The primary aim was to estimate the impact of any type of superficial, deep, and organ/space infections on the quality of life, as measured by utility score, of adults undergoing surgery.

**Methods**

**Inclusion and exclusion criteria**

The PICOS (Population, Intervention, Comparison, Outcomes, and Study) framework was used to guide the inclusion criteria. The population comprised adults (as specified by the trialists) receiving any surgery in any setting. Interventions included any adult with any confirmed SSI using any prespecified diagnostic criteria. Comparators were adults undergoing surgery with no SSI. The main outcome measure was utility as measured by a standardized, validated quality-of-life questionnaire, such as EuroQol 5D (EQ-5D®), EuroQol Group, Rotterdam. The Netherlands) or Short Form 36 (SF-36, Ware, Kosinski and Keller, Boston. USA®). Any study in which quality-of-life data were gathered prospectively or retrospectively from adults undergoing surgery, with such data presented for those with and without SSI at similar time points, was included. The main outcome effect measure was mean difference in EQ-5D® utility. Study designs that were anticipated to fulfill this criterion were RCTs, controlled trials, case-control studies, matched control studies, or economic evaluation or modelling studies that completed primary data collection for utility values. Information was sought directly from authors if missing from published reports.

Studies were excluded if no set diagnostic criteria for SSI were used, secondary quality-of-life data from published sources (for example, economic modelling studies using published utility values that they did not themselves gather) were used, and studies on children. Studies for which the full text was not available (either published, unpublished or directly from the authors) were also excluded.

**Search**

The systematic review was registered with PROSPERO (registration number CRD 42021262633). The registration acted as the protocol for this review. Studies that met the PICOS criteria and published in English were included. The following databases were searched from inception to April 2022: PubMed, MEDLINE via EBSCO host, CINAHL via EBSCO host, and the NHS Economic Evaluation Database.

A prespecified search strategy was formulated and adapted for each database’s conventions; details of the main search strategy are available in the supplementary material. The NHS Economic Evaluation Database was searched using Medical Subject Headings only. Reference lists in relevant articles were screened to identify any further potential papers.

**Study selection and data extraction**

Two authors independently examined the titles and abstracts identified by the search strategy to remove any duplicate records and irrelevant reports. Full-text versions of potentially relevant studies identified by at least one author were retrieved and evaluated. The same two authors independently assessed each study to determine whether it met the eligibility criteria, and extracted data using a standardized data extraction form developed for this review. Any disagreements were resolved by discussion between the authors, with a further author acting as arbiter. The data extraction form included the following general information (author(s), title, source, contact address, year of study, country of study, year of publication); trial characteristics (design, time horizon, quality-of-life instrument used); participants (baseline characteristics, type of surgery, inclusion and exclusion criteria, sample size, and number of patients allocated to each group or patient cohort details); interventions (SSI type; superficial, deep, organ/space); and outcomes (utility measure in both infected/non-infected groups).

All study outcomes were converted to a single index value of health state utility (EQ-5D®), if not already presented as such in the studies. SF-12® and SF-36® summary measures are not preference-based and as such cannot be used to subsequently calculate QALYs®. Conversion to EQ-5D® values is required, so SF-12® study-level data (mean and standard error) and SF-36® summary measures were converted to EQ-5D® preference scores using UK tariff values. SF-12® scores were converted using the two-variable model outlined by Lawrence and Fleishman. SF-36® data were converted using Model EQ (1) from Ara and Brazier. Synthesis was planned using a random-effects meta-analysis model, which assumes that the study effect sizes are different and that the collected studies represent a random sample from a larger population of studies. Heterogeneity was explored with forest plots and an $I^2$ test. In addition, the $P$ statistic describing the percentage variation that may be attributed to between-study heterogeneity was calculated. The prediction interval, used to estimate the true effect size and plotting a distribution of true effects, was also calculated. Finally, a funnel plot was generated. Subgroup analysis and meta-regression was planned by infection type according to Centers for Disease Control and Prevention criteria (superficial, deep, organ/space), wound classification (clean, clean contaminated, contaminated, dirty or infected), National Nosocomial Infections Surveillance risk index criteria, and time after surgery if possible. Analyses were undertaken using Stata® statistical software and the Prediction Intervals Program.

**Quality assessment**

Study quality assessment was completed independently by two authors using the STROBE checklist for case–control studies, the Drummond checklist for economic evaluations, and the Cochrane Risk of Bias (RoB 2) tool for RCTs. RCTs in which randomization of the intervention was not relevant to the SSI outcomes of primary interest of the present review were also assessed using the STROBE checklist. Study components were rated using the relevant ratings in each instrument. Utility valuation studies were rated against the components presented by Stalmeier et al. Any disagreement was resolved through discussion. Quality assessment was evaluated both within and across all included study types.

**Results**

**Study selection and characteristics**

The search identified 3142 titles and abstracts after removal of duplicates. A total of 43 articles were initially identified as having potential and full texts of these were retrieved. Seven articles reported SSI utility data included with the main study comparison groups. The authors were contacted directly for information, but none was obtained. These articles were subsequently excluded. A further 21 were excluded for other reasons and 15 studies met the inclusion criteria (Fig. 1). Characteristics of all included studies are summarized in Table S1.
The single economic evaluation\textsuperscript{28} used a case–control design, with the exception of two RCTs\textsuperscript{38,40} that provided a flow diagram. One\textsuperscript{42} did not address limitations, and another\textsuperscript{36} did not provide details of funding. All other studies performed well on checklist criteria relating to discussion and other information. Overall, the quality of the included studies was fair, with case–control methodology being associated with lower levels of reporting quality than economic evaluation and health state valuation. Results of the quality assessments for all included studies can be found in the Table S2 (case control studies), Table S3 (RCTs), Table S4 (economic evaluation), and Table S5 (health state valuation).

**Study-level utility impact of surgical-site infection**

Single-study estimates of the utility decrement associated with SSI ranged from \(-0.203\) in deep SSI following total joint replacement (34 cases of SSI) at more than 1 year after surgery\textsuperscript{41} to \(-0.03\) in all SSIs requiring antibiotic treatment 30 days after caesarean section (39 cases of SSI) using standard wound dressing\textsuperscript{44}. Table S6 outlines the study-level estimates of utility associated within the range of SSIs reported at various time points and SSI types.

**Meta-analyses**

The authors of studies that did not report the standard deviation, standard error or confidence interval of mean estimates for methods of the STROBE checklist\textsuperscript{28} were completed well. There was variability in checklist elements related to bias and statistical methods. Only two studies\textsuperscript{38,36} provided a flow diagram. One\textsuperscript{42} did not address limitations, and another\textsuperscript{36} did not provide details of funding. All other studies performed well on checklist criteria relating to discussion and other information. Overall, the quality of the included studies was fair, with case–control methodology being associated with lower levels of reporting quality than economic evaluation and health state valuation. Results of the quality assessments for all included studies can be found in the Table S2 (case control studies), Table S3 (RCTs), Table S4 (economic evaluation), and Table S5 (health state valuation).
quality-of-life scores were contacted directly by e-mail for missing data. One supplied the information, one was no longer available, and no response was received from the others. Thus, data from six studies across seven time points (1 to more than 12 months) and a range of procedures (arthroplasty, spinal, lower limb trauma, caesarean section, and all non-obstetric in 1 study site) were available for meta-analysis. The mean difference in EQ-5D™ utility in all studies combined was -0.08 (95 per cent c.i. -0.11 to -0.05; I² = 40 per cent) (Fig. 2).

**Subgroup analyses**

Data from two studies across five time points were available for meta-analysis of deep SSI only. The mean difference in EQ-5D™ utility was -0.11 (95 per cent c.i. -0.17 to -0.06).

---

**Fig. 2** Main meta-analysis outcome measure showing mean difference in EQ-5D™ utility estimates across six included studies

Mean differences are shown with 95% confidence intervals. A random-effects model was used for meta-analysis. *a-d* signify different time points within the same study or for Hyldig the two dressing comparator groups (i.e. both arms of the RCT).

**Fig. 3** Forest plot for EQ-5D™ mean difference in deep surgical-site infection only

Mean differences are shown with 95% confidence intervals. *Values are mean(s.d.): EQ-5D™ utility estimates. A random-effects REML model was used for meta-analysis. *a-d* signify different time points within the same study or for Hyldig the two dressing comparator groups (i.e. both arms of the RCT).
### Mean difference across all surgical-site infection types at different time points

<table>
<thead>
<tr>
<th>Reference</th>
<th>Mean difference</th>
<th>Mean difference</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>&lt; 4 months</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perencevich et al.</td>
<td>-0.11 (-0.17, -0.06)</td>
<td>12.94</td>
<td></td>
</tr>
<tr>
<td>Parker et al.</td>
<td>-0.10 (-0.23, 0.03)</td>
<td>3.98</td>
<td></td>
</tr>
<tr>
<td>Hyldig et al.</td>
<td>-0.13 (-0.17, -0.08)</td>
<td>14.72</td>
<td></td>
</tr>
<tr>
<td>Hyldig et al.</td>
<td>-0.03 (-0.07, 0.01)</td>
<td>17.86</td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: ( \tau^2 = 0.00 ), ( I^2 = 71.30% ), ( H^2 = 3.48 )</td>
<td>-0.09 (-0.14, -0.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test of ( \theta_1 = \theta_j ); Q(3) = 12.21, ( P = 0.01 )</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| **6–9 months**           |                 |                |            |
| Haddad et al.            | -0.11 (-0.20, -0.02) | 7.55          |
| Parker et al.            | -0.14 (-0.27, -0.01) | 4.13          |
| Parker et al.            | -0.07 (-0.21, 0.07) | 3.69          |
| Heterogeneity: \( \tau^2 = 0.00 \), \( I^2 = 0.02\% \), \( H^2 = 1.00 \) | -0.11 (-0.17, -0.04) |
| Test of \( \theta_1 = \theta_j \); Q(2) = 0.51, \( P = 0.78 \) |                 |                |            |

| **≥ 12 months**          |                 |                |            |
| Aboltins et al.          | -0.07 (-0.17, 0.04) | 6.18          |
| Guirro et al.            | -0.03 (-0.09, 0.03) | 12.11         |
| Haddad et al.            | -0.08 (-0.18, 0.01) | 6.87          |
| Haddad et al.            | -0.06 (-0.16, 0.04) | 6.33          |
| Parker et al.            | -0.16 (-0.30, -0.02) | 3.63         |
| Heterogeneity: \( \tau^2 = 0.00 \), \( I^2 = 0.02\% \), \( H^2 = 1.00 \) | -0.06 (-0.10, -0.02) |
| Test of \( \theta_1 = \theta_j \); Q(4) = 2.86, \( P = 0.58 \) |                 |                |            |

| **Overall**              |                 |                |            |
| Heterogeneity: \( \tau^2 = 0.00 \), \( I^2 = 39.72\% \), \( H^2 = 1.66 \) | -0.08 (-0.11, -0.05) |
| Test of \( \theta_1 = \theta_j \); Q(11) = 17.13, \( P = 0.10 \) |                 |                |            |
| Test of group differences: \( Q_b(2) = 1.73 \), \( P = 0.42 \) |                 |                |            |

**Fig. 4 Forest plot for EQ-5D™ mean difference across all surgical-site infection types at different time points**

Mean differences are shown with 95% confidence intervals. A random-effects REML model was used for meta-analysis. *a–d* signify different time points within the same study or for Hyldig the two dressing comparator groups (i.e. both arms of the RCT).

**Fig. 5 Funnel plot of all included studies**
utility associated with deep SSI was −0.10 (95 per cent c.i. −0.14 to −0.06; I² = 0 per cent) (Fig. 3).

To explore the impact of time since surgery, eligible studies were regrouped into three time-point categories of less than 4 months, 4–9 months, and 12 months or more. The mean difference in EQ-5D™ utility was −0.09 (−0.14 to −0.04; I² = 71 per cent) at less than 4 months, −0.11 (−0.17 to −0.04; I² = 0.02 per cent) at 6–9 months, and −0.06 (−0.10 to −0.02; I² = 0.02 per cent) at 12 months or more (Fig. 4). No data were available for the other planned subgroup analyses.

Heterogeneity
The estimate of mean difference in EQ-5D™ utility across all studies may represent moderate heterogeneity with an I² value of 40 per cent. An asymmetric funnel plot, as evidenced by scatter points outwith the pseudo 95 per cent confidence interval either side of the mean summary effect estimate, may indicate this is due to publication bias (Fig. 5). Lower I² values from subgroup analyses in which only one type of SSI was considered (for example I² = 0 per cent for deep SSI) and higher ones when different types of surgery were considered (for example I² = 71 per cent for all SSI types at less than 4 months since surgery) suggest a clinical source of the overall moderate heterogeneity found in the main meta-analysis outcome measure. For the main outcome measure, the true effect size in 95 per cent of all comparable populations fell in the interval −0.16 to −0.01 (Figure S6).

Discussion
The primary aim of the review was to estimate the impact of an SSI on quality of life, as measured by utility score, for adults undergoing surgery of any type for all SSIs.

In the meta-analysis of six studies across all time points since surgery, the mean difference in EQ-5D™ utility across all SSI types was −0.08 (95 per cent c.i. −0.11 to −0.05; I² = 40 per cent) (Fig. 2). For deep SSI, the utility decrement associated with these infections dropped to −0.10 (−0.14 to −0.06; I² = 0 per cent). The greatest utility decrement was associated with the interval 6–9 months after surgery (−0.11, −0.17 to −0.04; I² = 0.02 per cent) and this reduced with time (−0.06, −0.10 to −0.02; I² = 0.02 per cent, at 12 months or more).

The 0.08 utility decrement in the primary outcome meta-analysis may appear relatively minimal. There are, however, certain contextual and methodological considerations to acknowledge. Analysing beyond the EQ-5D™ overall mean score provides insight into which aspects of health are contributing to the reported changes. Thus, a 0.08 drop in overall EQ-5D™ utility may represent changes in up to two health state dimensions of the valuation instrument. In practical terms, this could mean that SSIs have influenced health to the extent that moderate, severe or even extreme problems are now present in mobility, self-care, usual activities, pain/discomfort or anxiety/depression. Indeed, given the prediction interval from the analysis, some patients may experience decrements of up to 0.16, compounding this effect. The distribution of true effects indicates that 95 per cent of patients in all comparable populations would consistently experience a decrement in utility due to SSI, highlighting the importance of this clinical issue. The utility value decrement must also be viewed in the context of the high global prevalence of SSI, resulting in significant cumulative effects on health at the population level. The finding that this decrement is present even at 12 months or more after surgery further highlights the importance of prevention efforts and investment in this area.

To the authors’ knowledge, this is the first systematic review to combine utility valuations of SSI in a meta-analysis. Previous work that aimed to summarize the evidence base on SSI utility values identified 28 studies, but no meta-analysis was completed. Additionally, only nine of these studies used patient-level data for health state valuations. The remaining 19 used values mainly sourced from the literature and were of poor methodological quality. The range of utility decrement reported is comparable at the lower end of estimates (0.04 versus 0.03 in this study) but not at high the higher end (0.48 versus 0.203 in this study). Interestingly, the authors concluded that the SSI utility decrement is suggested to be near 0.1, in alignment with the findings from the meta-analysis presented here, with an overall mean difference in EQ-5D™ utility of 0.08 in all studies combined.

The scope for heterogeneity across studies examining SSI is wide because of the diverse nature of the infection, evidenced by multilevel diagnostic criteria, surgical classifications, and specialties. The lack of previous meta-analyses in this field has been justified on this basis. The included studies for this review predominantly examined deep SSI in orthopaedic surgery and so meta-analysis was deemed appropriate. Studies were not excluded based on surgery and infection type, in line with the primary aim of the review. The moderate levels of heterogeneity found in the main meta-analysis outcome measure support this decision.

Overall, the methodological quality of the studies in this review was good. Several studies did not, however, report the standard deviation, standard error or confidence interval for mean estimates of quality-of-life scores, and only one author responded to the request for such information. Thus, the scope for meta-analysis was limited by the reporting in the original studies. This highlights the need for consistency in reporting of future work in this field, such as the use of reporting tools or checklists for peer review. A further limitation of this review is the lack of data relating to superficial and organ-space SSI and surgical classifications. Although the orthopaedic focus of most included studies facilitated meta-analysis, the utility decrements reported are heavily weighted towards this surgical specialty and may not be appropriate for use in other surgical settings.

Global guidelines on the prevention of SSI from the WHO specifically call for robust SSI economic and burden studies to address clinical need. The present study contributes to this call. As with European SSI surveillance, where specific surgical categories are used as indicator metrics for overall SSI burden, the present findings may offer insight into a wider scope of surgical specialty than those of the studies in the meta-analyses (orthopaedic and obstetrics). Given the paucity of synthesized estimates of SSI impact on health-related quality of life, the review is presented as a reference point from which clinicians may begin to infer utility decrements within wider specialty groupings based on associated clinical risk of other surgical categories. A synthesized estimate offers methodological advantage and increased precision over those derived from single studies. As such, the present review provides a higher rank of evidence for this purpose than previously published work.

The findings of this study confirm the significant burden that SSIs present to the surgical patient population both in the short and long term. The authors sought to quantify the known clinical burden of these infections and demonstrate the long-term (12 months after surgery and beyond) impact on patient quality of life. EQ-5D™ utility estimates for a range of SSIs for use in future economic modelling are also provided. Recommendations for future work in this area include improvement in the quality and consistency of reporting...
quality-of-life outcomes and data collection across the full scope of the SSI clinical pathway and timeline.

Funding
Funding for this work was received from Glasgow Caledonian University and Health Protection Scotland via the Evaluation of the Cost of Nosocomial Infection Study.

Author contributions
Agi McFarland (Conceptualization, Data curation, Formal analysis, Investigation, Project administration, Writing—original draft, Writing—review & editing), Sarkis Manoukian (Conceptualization, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Writing—review & editing), Helen Mason (Investigation, Resources, Supervision, Validation, Writing—review & editing), and Jacqui Reilly (Funding acquisition, Methodology, Resources, Supervision, Validation, Writing—review & editing).

Disclosure
The authors declare no conflict of interest.

Supplementary material
Supplementary material is available at B/ES online.

Data availability
Data generated for this review are available directly from the corresponding author.

References
21. Ara R, Brazier J. Deriving an algorithm to convert the eight mean SF-36 dimension scores into a mean EQ-5D preference-based score from published studies (where patient level data are not available). Value Health 2008; 11:1131–1143
26. Timberlake. Stata. 2022; 17
28. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: