Is administration of proton pump inhibitors in functional dyspepsia worth the risk of developing gastric cancer: a Markov model to bridge the gap between scientific evidence and clinical practice

Efrat Broide, Adi Eindor-Abarbanel, Haim Shirin, Vered Richter, Shay Matalon, Moshe Leshno

ABSTRACT

Objective To formulate a decision analysis model based on recently published data that addresses the dilemma, whether improvement in quality of life rationalises continued proton pump inhibitors (PPIs) use despite the risk of gastric cancer (GC) in patients with functional dyspepsia (FD).

Design A Markov model consisting of an initial decision regarding treatment and PPI (denoting it by PPI strategy) or any other treatment without PPI (denoting it by placebo strategy) was designed.

Data sources Data from prospective cross-sectional studies indicating risk stratification for GC after the use of PPI, combined with a Markov model that comprised the following states: Live, GC stages 1–4, Death.

Outcome measures The primary outputs included quality-adjusted life years (QALYs) and life expectancy (LE).

Sensitivity analyses were performed to evaluate the robustness of the model and address uncertainty in the estimation of model parameters.

Setting We considered only patients whose symptoms were relieved with PPIs and thus, had a better quality of life compared with patients who did not receive PPIs.

Results The base case model showed that PPIs compared with placebo decreased LE by 58.4 days with a gain of 2.1 QALY. If utility (quality of life of patients with FD using PPI compared with patients with FD without PPI) improved by more than 0.8%, PPI use is considered better than placebo. Older patients benefited less from PPI treatment than did younger patients.

Conclusion To bridge the gap between evidence and decision making, we found that even a small improvement in utility from the administration of PPIs at any time point.

Strengths and limitations of this study

- A comprehensive simulation that includes all possible health states that could be experienced by patients with functional dyspepsia responding to proton pump inhibitor (PPI) treatment.
- Formulation of an analytical model which can assess the risk of developing gastric cancer versus the utility from the administration of PPIs at any time point.
- The cost-effectiveness analysis makes it easier to understand the potential economic benefits of adopting the ideal strategy.
- Certain assumptions are made which may affect the results.
- The assumptions made for this study, are not necessarily generalisable to individuals on PPIs for other indications, or outside the age window investigated.

INTRODUCTION

Proton pump inhibitors (PPIs) have been available in USA since the mid-1980s. They are commonly prescribed to prevent gastro-oesophageal reflux disease (GORD) and peptic ulcer disease as well as for eradication of Helicobacter pylori and improvement of non-ulcer dyspepsia.

Over the last few decades, PPIs have also been used as maintenance therapy to reduce the risk of peptic ulcer disease associated with non-steroidal anti-inflammatory drug use and low-dose aspirin therapy for prolonged periods.

With the widespread use of PPIs, the issue of long-term safety was raised. Some adverse events were reported in patients with GORD who used PPIs for long periods. The most common include bone fractures, enteric infections, (specifically Clostridium difficile) community-acquired pneumonia (with pooled ORs of 1.44–2.65, 1.69–3.3 and 1.04–1.92, respectively) and deficiencies in nutrients such as iron, vitamin B<sub>12</sub>, and calcium.
and chronic kidney disease. The putative therapeutic effect of PPIs is direct blockade of the potassium influx and hydrogen outflux, thus inhibiting the acid production by parietal cells. As a consequence, atrophic gastritis, a premorbid condition of gastric adenocarcinoma, could worsen. Concerns have been raised about the potentially increased risk of gastric cancer (GC) following long-term PPI use. Few epidemiological studies have evaluated the association between PPI use and GC risk. A recent meta-analysis showed that the risk of GC is increased by 43% among PPI users, but these studies included both *H. pylori*-infected and non-infected subjects. Therefore, it is difficult to conclude whether the use of PPIs or the persistent infection was the cause of this increase. Two Danish studies noted an association between GC and PPI use. However, as PPIs are given to symptomatic patients, this association may also be a result of a selection bias. Two recent studies, based on a large health database in Hong Kong and Sweden, reported an even higher risk for GC among long-term PPI users, with an HR of 2.44 and Standardised incidence ratio (SIR) of 3.38, respectively. This risk increased with the duration of PPI use and was documented even after *H. pylori* eradication. Note that although the HR of developing GC with PPI is high, the absolute risk of GC under treatment with PPI remains low.

Functional dyspepsia (FD), which affects approximately 10% of the adult population, is a disorder that can markedly impair quality of life. This syndrome can be divided into epigastric pain syndrome (EPS) (18% of patients), postprandial distress syndrome (61%) and overlapping syndrome (21%). PPIs were shown to be more effective than placebo for improving global symptoms and quality of life in people with FD-EPS. Therefore, they are considered a therapeutic option for this population. Assuming that long-term use of PPIs in patients with FD may increase the risk of GC on the one hand, but improve quality of life on the other, a decision making tool to evaluate the harm versus benefit in these patients is required.

We need to use models to bridge the gap between primary evidence and guideline development. The models are usually mathematical frameworks that play an important role in integrating and extending the evidence on outcomes of healthcare interventions.

The objective of this study was to formulate a decision analysis model, based on recently published data, to address the question of whether improvement in quality of life rationalises continuity of PPI treatment despite the risk of GC. Moreover, we believe that based on our results, we will be able to extrapolate the use of PPIs to other, more common, indications such as gastro-oesophageal reflux, in which the benefit of using PPIs is even higher.

**METHODS**

Model structure

The structure of the Markov model (a mathematical modeling technique that is used to describe the transitions a cohort (or Monte Carlo simulation) of patients make among a number of mutually exclusive and exhaustive health states during a series of given intervals) consisted of an initial decision regarding treatment with PPI or any other treatment for dyspepsia without PPI (eg, histamine-2 receptor antagonists). We considered only patients whose symptoms were relieved with PPIs and thus, had a better quality of life compared with patients who did not receive PPIs. The model is presented graphically in figure 1. The sources for the numbers in the model were extracted from published articles listed in table 1. No specific permissions were required to access the data, since we used only published data.

The initial health state in the Markov model is called ‘Live’. All patients, whether they did or did not initially receive PPI, start in this health state (‘Live’) (figure 1). Patients will remain in this health state until either developing GC or dying from other reasons. The mortality rate in this health state is based on the US life table of the general population (a table which shows, for each age, what is the probability that a person of that age will die). Every patient may develop GC, but the rate of developing GC in patients who received PPI was greater (HR 3–10) than in patients who did not receive PPI. The annual incidence of GC per 100 000 people is 5–7. In the model we assumed that the incidence is 10 per 100 000, which is higher than published data and favours not using PPI. Patients who developed GC were randomly stratified by cancer stages, where each stage is a health state in the Markov model. We assume that the distribution of the stratification to cancer stages of people who developed GC is the same for patients who received PPI and for patients who did not receive PPI. The distribution of the

---

**Figure 1** The structure of the Markov model consisted of an initial decision regarding treatment with proton pump inhibitor (PPI) or any other treatment for dyspepsia without PPI (eg, histamine-2 receptor antagonists). Patients who initially received PPI or those who did not could have developed gastric cancer (GC), but the rate of developing GC in patients who received PPI was greater (HR 3–10).

---

Broide E, et al. BMJ Open 2020;10:e031091. doi:10.1136/bmjopen-2019-031091
Table 1  Parameters in the model

| Description (variables) | Baseline value | Range       |
|-------------------------|----------------|-------------|
| **Weibull parameters estimation of survival curves** |               |             |
| GC stage I              |               |             |
| γparameter for Weibull distribution of survival in stage I GC | 0.7322 | 0.71   | 0.99   |
| λparameter for Weibull distribution of survival in stage I GC | 0.0822 | 0.05   | 0.12   |
| GC stage II             |               |             |
| γparameter for Weibull distribution of survival in stage II GC | 0.71519 | 0.65  | 0.75   |
| λparameter for Weibull distribution of survival in stage II GC | 0.2331 | 0.15  | 0.30   |
| GC stage III            |               |             |
| γparameter for Weibull distribution of survival in stage III GC | 0.77248 | 0.70  | 1.25   |
| λparameter for Weibull distribution of survival in stage III GC | 0.4527 | 0.30  | 0.55   |
| GC stage IV             |               |             |
| γparameter for Weibull distribution of survival in stage IV GC | 0.82569 | 0.75  | 0.88   |
| λparameter for Weibull distribution of survival in stage IV GC | 0.7452 | 0.60  | 0.80   |
| **Distribution of GC stages** |               |             |
| Proportion of GC diagnosed in stage I \(^{(18)}\) | 0.102 | 0.08  | 0.12   |
| Proportion of GC diagnosed in stage II \(^{(18)}\) | 0.286 | 0.25  | 0.3    |
| Proportion of GC diagnosed in stage III \(^{(18)}\) | 0.441 | 0.4   | 0.45   |
| **Utility values**      |               |             |
| Percentage increased utility with PPI | 0.1     | 0  | 0.5    |
| Utility of cancer \(^{(22)}\) | 0.5     | 0.4  | 0.75   |
| Utility of dyspepsia \(^{(26)}\) | 0.8     | 0  | 1      |
| **General assumptions** |               |             |
| Age, years              | 50            | 50  | 70     |
| *HR for GC with PPI \(^{(13,14)}\) | 5.04 | 1  | 34     |
| Rate of GC without PPI \(^{(15)}\) | 0.0001 | 0.00001 | 0.00002 |
| Time horizon            | 50            | 20  | 60     |

*The basic HR for GC with PPI was 5.04 for the first year, 6.65 for the second year and 8.34 for the third year or more.
GC, gastric cancer; PPI, proton pump inhibitor.

The stage of the GC is based on the American Joint Committee of Cancer (AJCC) Cancer Staging Manual (2017)\(^{(21)}\) and is presented in Table 1.

The primary outputs of the model included quality-adjusted life years (QALYs) and life expectancy (LE). The model cycle length is 1 year.

The Markov model was implemented with TreeAge Pro 2019 software (https://www.treeage.com). Statistical and Weibull fitting analyses were performed using MATLAB 2018 software.

Model survival estimates

The overall mortality rate, which corresponded to the probability of death, was derived from the Kaplan-Meier curves for overall survival (OS) among patients with GC who underwent resection and received chemotherapy and/or radiation prior to surgery, stratified by prognostic stage group (National Cancer Database, eighth edition, 2017). GetData Graph Digitizer software (http://getdata-graph-digitizer.com) was used to extract the data points from the OS curves. These data points were then used to fit parametrical survival models. Data points were applied of its parametrical survival models. Weibull, log-logistic and log-normal distribution were chosen to simulate Progression Free Survival (PFS) and OS curves based on goodness of fit to the follow-up survival data. We used the Nelder-Mead algorithm to fit all OS curves to: Weibull, log-logistic and log-normal distributions. The transition probabilities (tp) in the model were computed by: 

\[
\text{tp} = 1 - \frac{S(t+z)}{S(t)}
\]

where \(S(t)\) is the cumulative survival curve (Pr \((T > t)\), \(T\) is the distribution of the events along time, Pr stands for probability and \(z\) is the cycle length (in our case it is 1 year). The mathematical proof is easy and is given by:

\[
\Pr(T \leq t + z | T > t) = 1 - \Pr(T > t + z | T > t) = 1 - \frac{S(t+z)}{S(t)}.
\]

In this study for the Weibull distribution, 

\[
S(t) = \exp(-\lambda t^\gamma)
\]

where \(\lambda\) and \(\gamma\) are the parameters of the Weibull distribution, for the log-logistic distribution,

\[
S(t) = \frac{1}{2} - \frac{1}{2} \text{erf}\left(\frac{\ln(t) - \mu}{\sqrt{2} \sigma}\right)
\]

and for the log-normal distribution,
Utility estimates
The health states of GC were assigned a health utility score based on quality of life data in the Tufts Medical Center Cost effectiveness analysis (CEA) Registry database (http://healtheconomics.tuftsmedicalcenter.org). The utility value related to the quality of life regarding dyspepsia without PPI and improvement in quality of life with PPI was tested with sensitivity analyses over a large range of values (patient preferences). We used the Tufts Medical Center CEA Registry under the term ‘dyspepsia’ to find the ‘Utility Weights’ of dyspepsia. As the values in different papers range between 0.7 and 0.97, we used 0.8 utility value with the largest range (between 0 to 1).

The list of all the parameters in the model with their values and range for one-way sensitivity analysis is presented in table 1.

To estimate the efficacy (QALY) and LE, we used Monte Carlo simulation of 100 samples where each sample included 100 000 subjects. Because the rate of mortality of GC depends on the duration of the disease, cohort (the Markov cohort model is a smooth model of the proportion of a cohort in each state at each time. Monte Carlo simulation or microsimulation simulates many individuals) estimation was not applicable (The authors thank Stuart Mealing for pointing it).

In addition, we used microsimulation of 1000 samples where each sample included 1000 subjects to display the distributions of the QALY and LE outcomes of the 1000 trials comparing the PPI strategy to the placebo strategy.

Sensitivity analysis
We performed internal model validations, which demonstrated that the OS curves generated by the Weibull fitting closely approximated those presented in the real data of the Kaplan-Meier OS curves (figure 2). A series of sensitivity analyses was performed to evaluate the robustness of the model and address uncertainty in estimation of model parameters. Utilities were varied over large value intervals. In univariable sensitivity analyses, we varied the value of one parameter at a time over its range. The baseline values and ranges of the model’s parameters are listed in table 1.

Patient and public involvement
Patients and the public were not involved in the design or planning of the study.

RESULTS
Base case results
The best fitting OS curve was the log-logistic distribution and therefore we present the results based on log-logistic distribution fitting. Note that the results of other distributions were very similar to the log-logistic distribution. The OS curves in each GC pathological stage based on the National Cancer Database and the log-logistic fitting, respectively, are shown in figure 2. The base case model results showed that the LE was 29.81±0.12 and 29.83±0.12 years in the placebo and PPI strategies, respectively. Thus, PPI compared with placebo resulted in a decrease of 7.3 days in LE (0.02 years). However, the QALY was 26.2±0.03 and 24.04±0.03 in the placebo and PPI strategies, respectively. PPI is more effective than placebo with a 2.16 QALY gain (table 2). The distributions of QALY and LE of a microsimulation of 1000 samples (each sample included 1000 subjects) are given in figure 3, comparing the PPI strategy to the placebo strategy.

Sensitivity analyses
The results of univariable sensitivity analyses are presented in a Tornado diagram (the Tornado diagram is a set of one-way sensitivity analyses brought together in a single bar chart of all the variables in the model). The variables are listed vertically, and the influence of the variables on the QALY gain are ordered so that the largest bar appears (largest QALY impact) at the top of the chart, the second largest appears second from the top, and so on (figure 4). The most influential variable was: percentage of increasing utility with PPI. This was followed by the age of the patients and time horizon. For an increase in improvement utility value with PPI of more than 0.8% than the utility value of placebo, the use of PPI was better than placebo strategy (figure 5). The benefit of PPI was less for older patients than it was for younger patients. In two-way sensitivity analysis with the variables percentage of increasing utility with PPI and HR of GC after 3 years of PPI use (range 8–34), only the first was relevant (figure 6).
**DISCUSSION**

In the era of evidence-based medicine, when a tremendous amount of scientific data are published, it is often difficult to bridge the gap between newly published data and routine, practical clinical guidelines. Therefore, it is important to find a tool that enables implementing robust scientific evidence into new strategies and interventions. Although the two largest population-based cohort studies that suggested that long-term PPI use is associated with an increased risk of GC \cite{13,14} had some limitations, they are still the most comprehensive studies published so far. These articles triggered panic among clinicians prescribing PPIs and among patients consuming the drugs. Quantitative methods for evaluating decisions between multiple alternatives in situations of uncertainty (meaning translating newly published data (evidence) into a simple everyday decision making tool) are needed. Decision analysis models are effective for objective evaluation of outcomes in specific clinical settings.

The Markov model is a mathematical modelling technique that describes the transitions a cohort of patients makes among a number of mutually exclusive and exhaustive health states, during a series of short intervals or cycles. \cite{22} Markov models are particularly useful when a decision problem involves an ongoing risk over time. \cite{23} We used this model to clarify the dynamics and trade-offs involved in selecting the best expected outcome.

Our model focused on the controversial clinical group of patients with FD treated with PPIs, who constitute approximately 10% of the adult population. \cite{12} These patients often experience troublesome symptoms and need PPI treatment for extended periods to improve their quality of life. \cite{24} The Markov model used in the current study is a personalised calculator for GC risk in such patients. As QALY is the most important issue for patients with FD, \cite{25} we chose this factor as the leading favourable outcome in our study. Even though we used the most extreme HR values for developing GC, PPI treatment was calculated as preferred over placebo, even when the QALY improved by only 0.8%.

### Table 2

| Strategy   | QALY   | Incremental QALY | LE (years) | Incremental LE (years) |
|------------|--------|------------------|-----------|------------------------|
| Placebo    | 24.04  | –                | 29.81     | 0.02                   |
| PPI        | 26.20  | 2.16             | 29.83     | –                      |

PPI, proton pump inhibitor.

**Figure 3** The distributions of quality-adjusted life years (QALYs) and life expectancy (LE) of a microsimulation of 1000 samples (each sample included 1000 subjects), comparing the proton pump inhibitor (PPI) strategy to the placebo strategy.

**Figure 4** Tornado diagram. The most influential variable was: percentage of increasing utility with proton pump inhibitor (PPI). This was followed by the age of the patients and time horizon. GC, gastric cancer.
Sometimes a model can be made more accurate but at the expense of simplicity. In cases like this, the simpler model may actually be superior, because it enables visualisation of a process, better understanding of it and easier decision making. Similarly, the Markov model in our study does not present the whole real world practice, but focuses on a specific harmful result of using PPI (GC). In addition, the current analysis was limited by our assumptions. A key issue in the model was to select the appropriate ‘benefit from PPI’ arm of the model. In the sensitivity analysis, we used the most extreme low values from the published range of data to account for possible inaccuracies. Nevertheless, the comparison between improvement in quality of life and the burden of GC is not equal in all people. Furthermore, it is important to note that although the HR of developing GC with PPI is high, the absolute risk of GC under treatment with PPI remains low. Additionally, the findings of this study are not necessarily generalisable to individuals on PPIs for other indications, or outside the age-window investigated. Also, we used data that were investigated on several populations but extrapolation to all ethnicities cannot be taken for granted.

Therefore, we advise using our calculator (can be available by request) to obtain a more personalised and accurate decision for patients with FD. A study like this can contribute to the conversation with a patient around the possibility of continuing or discontinuing PPI use to weigh up the benefit of PPI use versus the risk of malignancy. But it must be understood that the power of this tool is in pointing out the validity of proposed treatment options, rather than a definitive decision making tool. In addition, although patients with GORD and peptic ulcer disease are likely to benefit more from PPI usage, a future analysis should be done in order to investigate it.

In conclusion, in our attempt to bridge the gap between evidence and decision making, we found that it is important to discontinue the use of PPIs when symptoms do not improve. However, when there is even minimal improvement in the patient’s quality of life, continued PPI treatment is justified.

Author affiliations
1The Kamila Gonczarowski Institute of Gastroenterology and Liver Diseases, Shamir Medical Center (Assaf Harofeh Medical), Zerifin, Israel
2Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel
3Coller School of Management, Tel Aviv University, Tel Aviv, Israel

Contributors EB, AE-A, ML performed the research and analysed the data. EB, AE-A, HS, SM, VR collected the data. EB, AE-A, ML designed the research study and wrote the paper. HS, SM, VR contributed to the study design.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially.
and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID ID**
Adi Eindor-Ababaranel http://orcid.org/0000-0001-9677-0453

**REFERENCES**

1. Shi S, Klotz U. Proton pump inhibitors: an update of their clinical use and pharmacokinetics. *Eur J Clin Pharmacol* 2008;64:935–51.

2. Abramowitz J, Thakkar P, Isa A, et al. Proton pump inhibitor therapy: an overview of systematic reviews. *Otolaryngol Head Neck Surg* 2016;155:547–54.

3. Ali T, Roberts DN, Tierney WM. Long-term safety concerns with proton pump inhibitors. *Am J Med* 2009.

4. Scheiman JM. The use of proton pump inhibitors in treating and preventing NSAID-induced mucosal damage. *Arthritis Res Ther* 2013.

5. Lazarus B, Chen Y, Wilson FP, et al. Proton pump inhibitor use and the risk of chronic kidney disease. *JAMA Intern Med* 2016.

6. Nochaiswong S, Ruengorn C, Awiphon R, et al. The association between proton pump inhibitor use and the risk of adverse kidney outcomes: a systematic review and meta-analysis. *Nephrol Dial Transplant* 2018.

7. Grant CH, Gillis KA, Lees JS, et al. Proton pump inhibitor use and progression to major adverse renal events: a competing risk analysis. *QJM An Int J Med* 2019.

8. Poulsen AH, Christensen S, McLaughlin JK, et al. Proton pump inhibitors and risk of gastric cancer: a population-based cohort study. *Br J Cancer* 2009;100:1503–7.

9. Tran-Duy A, Spaetgens B, Hoes AW, et al. Use of proton pump inhibitors and risks of fundic gland polyps and gastric cancer: systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2016;14:1706–19.

10. Hvid-Jensen F, Pedersen L, Funch-Jensen P, et al. Maintenance proton pump inhibitor therapy with omeprazole and risk of gastric cancer: a population-based cohort study in the UK. *Gut* 2006;55:1538–44.

11. Vigneri S, Termini R, Leandro G, et al. A prospective randomized controlled trial using pantoprazole. *BMC Gastroenterol* 2011;11.

12. Moayyedi PM, Lacey BE, Andrews CN, et al. ACG and CAG clinical guideline: management of dyspepsia. *Am J Gastroenterol* 2017;112:988–1013.

13. Baghossian TA, Rashid FJ, Thompson W, et al. Deprescribing versus continuation of chronic proton pump inhibitor use in adults. *Cochrane Database Syst Rev* 2017;16.

14. Smith PM, Kerr GD, Cockel R, et al. Use of proton pump inhibitors and risk of gastric cancers: review of literature. *Best Pract Res Clin Gastroenterol* 2017;112:988–1013.

15. Dong J, Thrift AP. Alcohol, smoking and risk of oesophago-gastric cancer. *Best Pract Res Clin Gastroenterol* 2017;31:509–17.

16. Yamawaki H, Futagami S, Wakabayashi M, et al. Management of functional dyspepsia: state of the art and emerging therapies. *Ther Adv Chronic Dis* 2018;9:23–32.