

## RESEARCH

# Influence of experience on performance of individual surgeons in thyroid surgery: prospective cross sectional multicentre study

 OPEN ACCESS

Antoine Duclos *assistant professor of public health*<sup>1,2,11</sup>, Jean-Louis Peix *professor of surgery*<sup>3</sup>, Cyrille Colin *professor of public health*<sup>1,2</sup>, Jean-Louis Kraimps *professor of surgery*<sup>4</sup>, Fabrice Menegaux *professor of surgery*<sup>5</sup>, François Pattou *professor of surgery*<sup>6,7</sup>, Frédéric Sebag *professor of surgery*<sup>8</sup>, Sandrine Touzet *senior epidemiologist*<sup>1,2</sup>, Stéphanie Bourdy *project manager*<sup>1,2</sup>, Nicolas Voirin *biostatistician*<sup>9,10</sup>, Jean-Christophe Lifante *professor of surgery*<sup>3</sup>, The CATHY Study Group\*

<sup>1</sup>Hospices Civils de Lyon, Pôle Information Médicale Evaluation Recherche, Lyon F-69003, France; <sup>2</sup>Université de Lyon, Equipe d'Accueil Mixte 4128 Santé-Individu-Société, Lyon F-69002; <sup>3</sup>Hospices Civils de Lyon, Centre Hospitalier Lyon Sud, Service de Chirurgie Générale et Endocrinienne, Pierre Bénite, Lyon; <sup>4</sup>Department of Endocrine Surgery, Poitiers University, Jean Bernard Hospital, Poitiers, France; <sup>5</sup>Assistance Publique-Hôpitaux de Paris, Hôpital la Pitié-Salpêtrière, Service de Chirurgie Générale, Viscérale et Endocrinienne, Paris, France; <sup>6</sup>Centre Hospitalier Régional Universitaire de Lille, Chirurgie Générale et Endocrinienne, Lille, France; <sup>7</sup>Université Lille Nord de France, Institut National de la Santé et de la Recherche Médicale (INSERM), Lille; <sup>8</sup>Assistance Publique-Hôpitaux de Marseille, Centre Hospitalier Universitaire la Timone-Adulte, Marseille, France; <sup>9</sup>Hospices Civils de Lyon, Hôpital Edouard Herriot, Service d'Hygiène, Epidémiologie et Prévention, Lyon; <sup>10</sup>Centre National de la Recherche Scientifique (CNRS), Unite Mixte de Recherche 5558, Laboratoire de Biométrie et Biologie Evolutive, Université de Lyon, Lyon; <sup>11</sup>Center for Surgery and Public Health, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

## Abstract

**Objective** To determine the association between surgeons' experience and postoperative complications in thyroid surgery.

**Design** Prospective cross sectional multicentre study.

**Setting** High volume referral centres in five academic hospitals in France.

**Participants** All patients who underwent a thyroidectomy undertaken by every surgeon in these hospitals from 1 April 2008 to 31 December 2009.

**Main outcome measures** Presence of two permanent major complications (recurrent laryngeal nerve palsy or hypoparathyroidism), six months after thyroid surgery. We used mixed effects logistic regression to determine the association between length of experience and postoperative complications.

**Results** 28 surgeons completed 3574 thyroid procedures during a one year period. Overall rates of recurrent laryngeal nerve palsy and hypoparathyroidism were 2.08% (95% confidence interval 1.53% to 2.67%) and 2.69% (2.10% to 3.31%), respectively. In a multivariate analysis, 20 years or more of practice was associated with increased probability of both recurrent laryngeal nerve palsy (odds ratio 3.06 (1.07 to 8.80),  $P=0.04$ ) and hypoparathyroidism (7.56 (1.79 to 31.99),  $P=0.01$ ).

Surgeons' performance had a concave association with their length of experience ( $P=0.036$ ) and age ( $P=0.035$ ); surgeons aged 35 to 50 years had better outcomes than their younger and older colleagues.

**Conclusions** Optimum individual performance in thyroid surgery cannot be passively achieved or maintained by accumulating experience. Factors contributing to poor performance in very experienced surgeons should be explored further.

## Introduction

The operating room is a high risk environment because of the volume<sup>1</sup> and complexity of procedures undertaken.<sup>2</sup> Adverse events during hospitalisation are often linked to surgical care,<sup>3</sup> with wide variation in mortality rates across hospitals.<sup>4</sup> Although most efforts to improve surgical care focus on technological advances,<sup>5</sup> it seems unlikely that the use of robotics or attempts to standardise work procedures to improve efficiency would erase errors and costs completely.<sup>6</sup> The impact of human factors on patient safety is crucial, and the surgeon's performance is a core element of successful surgery.

Correspondence to: A Duclos [antoineduclos@yahoo.fr](mailto:antoineduclos@yahoo.fr)

Extra material supplied by the author (see <http://www.bmj.com/content/344/bmj.d8041?tab=related#webextra>)

**Web appendix:** R code used to fit linear model with quadratic term for figure 3

Professional expertise results from a gradual improvement in performance within a specialty. Typically, experts reach their peak performance in various domains at age 30 to 50 years—that is, after about 10 years' experience in their specialty.<sup>7</sup> Older doctors who have been in practice for a long time after graduation might have less factual knowledge and be less likely to adhere to evidence based medicine,<sup>8</sup> which risks the safety of care. However, few studies have explored the association between clinicians' experience and performance using objective outcome measurements.

In surgery, performance can be assessed by the occurrence of postoperative complications, and experience can be measured by the surgeon's age or time spent in practice. The importance of the education and training used to gain experience among young surgeons is obvious.<sup>9</sup> Acquiring the necessary technical background can take time before individuals are able to master routine procedures and achieve better outcomes.<sup>10 11</sup> Although complication rates can vary greatly during a surgeon's career, the potential decline in performance among very experienced surgeons remains unclear. Previous studies have used retrospective analyses of hospital databases without considering the clustered nature of the data.<sup>12 13</sup> We aimed to model the association between the experience and performance of surgeons on postoperative complications in thyroid surgery.

## Methods

### Study design and population

We conducted a prospective cross sectional study between 1 April 2008 and 31 December 2009, in five academic hospitals (centres A-E). The participating hospitals were high volume referral centres in endocrine surgery, each undertaking at least 500 thyroidectomies a year. All 28 endocrine surgeons who did thyroid surgery in these centres participated in the study. Patient recruitment for one year after surgery began in April 2008 in centres A to C, May 2008 in centre D, and July 2008 in centre E. All patients who underwent a thyroid procedure were eligible for inclusion.

### Outcome measures

We considered two major complications of thyroid surgery: permanent recurrent laryngeal nerve palsy<sup>14</sup> and hypoparathyroidism. Based on the available evidence, participating centres used an active follow-up protocol.<sup>15 16</sup> Postoperative outcomes were systematically assessed during hospitalisation within 48 hours after the thyroid procedure. A second assessment was planned at least six months after the thyroidectomy to diagnose permanent complications.

The proportion of recurrent laryngeal nerve palsy was measured in patients who underwent unilateral or bilateral thyroid procedures. Postoperative vocal cord mobility was assessed by laryngoscopy.<sup>17</sup> Exclusion criteria for recurrent laryngeal nerve palsy analysis included pre-existing nerve palsy before the intervention, previous thyroid surgery with unknown pre-existing nerve palsy status, voluntary resection of nerves during intervention due to invasive carcinoma, and patient death during follow-up.

The proportion of hypoparathyroidism was measured in patients who underwent a bilateral thyroidectomy. We defined postoperative hypoparathyroidism as a serum calcium concentration of lower than 2 mmol/L or a requirement for vitamin D or calcium supplementation (or both) to maintain healthy calcium concentrations after thyroidectomy.<sup>18</sup> Exclusion criteria for hypoparathyroidism analysis included unilateral

lobectomy, pre-existing hypocalcaemia or substitutive treatment with calcium or vitamin D (or both) before intervention, previous thyroid surgery with unknown pre-existing hypoparathyroidism status, and patient death during follow-up.

Owing to organisational constraints at centre E, systematic laryngoscopy was not feasible after each thyroid procedure. Therefore, vocal cord mobility was objectively assessed by laryngoscopy only in instances of postoperative voice alterations. For this reason, we did not include centre E in the analysis for recurrent laryngeal nerve palsy.

### Data collection

Local investigators and staff received handouts and training to become familiar with data collection and procedures for monitoring surgical complications. After each thyroidectomy, the surgeon would systematically complete a patient report form, which included items regarding surgical indication and procedures, and the number of interventions that the surgeon had undertaken that day. Clinical research assistants completed data collection using medical records. These data included patient demographics and information on previous thyroid surgeries, thyroid specimen weights, postoperative supplementation with calcium or vitamin D (or both), calcium assay values, and assessments of vocal cord mobility.

Surgeons also completed individual questionnaires regarding their demographics, previous background, and professional experience, using a one off survey. The surgeon's length of experience was measured as the number of years spent in practice since graduation (that is, the end of residency). Before the analysis, we separated the experience variable into four categories to depict the successive steps in a surgeon's career in France: less than 2 years (that is, a very young surgeon starting a surgical fellowship), 2 to 4 years (junior surgeon ending a surgical fellowship), 5 to 19 years (senior surgeon), and 20 years and over (very experienced surgeon at the head of a surgical department).

We obtained the structural characteristics of each participating hospital from the French annual statistics on healthcare facilities in 2008.<sup>19</sup> We calculated the proportion of cases included in the study as the ratio of the number of thyroidectomies with a fully completed patient form to the number of eligible thyroidectomies recorded in the hospital administrative databases.

### Ethical considerations

The Research Committee for the Protection of Persons allowed the study in accordance with ethical directives. The National Advisory Committee on Information Processing in Material Research in the Field of Health also approved the study, regarding the anonymous processing of personal health information. The participating centres approved the study protocol without giving incentives to surgeons for their participation. Before surgery, patients received written information about personal data use, and gave verbal consent for sharing their data.

### Statistical analysis

We described the characteristics of the hospitals and surgeons using absolute frequencies with percentages for categorical variables. We calculated mean values with standard deviations and median values with minimum-maximum intervals for continuous variables. We calculated 95% confidence intervals of complication rates using an exact method based on binomial distribution and used the Spearman's rank correlation test to investigate correlation between surgeon's length of experience

and age, as well as the overall number of thyroidectomies previously undertaken. To identify variables associated with recurrent laryngeal nerve palsy and hypoparathyroidism, we compared thyroid procedures with and without complications. We used  $\chi^2$  and Mann-Whitney tests to compare categorical and continuous variables, respectively.

We identified the factors associated with the probability of each complication by calculating crude and adjusted odds ratios using a mixed effects logistic regression. This approach was appropriate for the hierarchical design of this study, in which patients were nested in surgeons, who in turn were nested in hospitals. In the models, we classified a patient's case mix as a fixed effect (that is, the patient's age, sex, body mass index, weight of thyroid specimen, and complexity of surgical case), the surgeon as both random and fixed effects (that is, the length of experience and number of surgical procedures done by the surgeon on the same day), and the hospital as a random effect.

Since patients with high risk of complications might have been cared for by more experienced surgeons, we adjusted outcomes for the complexity of surgical case. We based complexity measurement on patients' variables that were significantly associated with complications in bivariate analyses.

Consequently, we defined a complex surgical case as a procedure involving a patient with Graves' disease or a malignant neoplasm, an extended thyroidectomy, or a thyroidectomy with lymph node dissection. Model estimates and their 95% confidence intervals were obtained by using the `glmmPQL` function<sup>20</sup> of the R free software, which fits generalised linear mixed models by an approximate method for facilitating inference.<sup>21</sup> We also calculated a measure of the explained variation in final multivariate models using a variant of  $R^2$ , based on absolute deviations.<sup>22,23</sup> Percentages of total variance attributable to the patient, surgeon, and hospital centre levels were calculated by dividing  $R^2$  of a model with the level of interest by the  $R^2$  the final model with all levels included.

Furthermore, we plotted individual surgeons' performances according to their age or length of experience. Surgeons who operated on fewer than 30 patients were not considered in this performance curve analysis. Accordingly, the individual performances of 15 and 22 surgeons were plotted for recurrent laryngeal nerve palsy and hypoparathyroidism analyses, respectively. For each surgical case, we calculated the predicted probability of complication using parameter estimates from a multivariate logistic regression. The variables used were identical to those used previously for assessing the risk of complications, except for the experience of the surgeon, which was the variable of interest.

For each surgeon, we calculated the expected and observed number of complications as the sum of the predicted complication probabilities and the sum of the observed complications for all the procedures that each surgeon had undertaken, respectively. We measured performance as the difference between the expected and observed complication rates, using the equation  $(E-O/\text{number of procedures}) \times 100$ , where E and O are the expected and observed number of complications, respectively. A value above zero indicated good performance, whereas a value below zero indicated poor performance.

We tested whether a trend existed in the performance curves by adjusting a linear model with experience entered both as a linear and a quadratic term. A significant quadratic term suggested that surgeons' performances did not vary linearly. Instead, performance might have increased, reached a peak, and then

decreased with age or length of experience. All tests were two tailed, and  $P < 0.05$  was considered significant.

## Results

Of 3679 eligible procedures, 3574 (97%) were completed during the study period. In accordance with the exclusion criteria and complete follow-up, 2357 (66%) procedures were selected for the analysis of recurrent laryngeal nerve palsy (fig 1), and 2904 (81%) for the analysis of hypoparathyroidism (fig 2).

We found no significant difference in case mix between patients who were included in analysis and those who were lost to follow-up. Overall rates of permanent nerve palsy and hypoparathyroidism were 2.08% (49/2357, 95% confidence interval 1.53% to 2.67%) and 2.69% (78/2904, 2.10% to 3.31%), respectively. We did not record any bilateral injuries of recurrent nerves.

Table 1 summarises characteristics of the 28 participating surgeons and their distribution among the centres. The surgeons' mean age was 41 years, with a mean length of experience of 10 years. Length of experience correlated strongly with age ( $\rho = 0.94$ ,  $P < 0.0001$ ) as well as with the overall number of thyroidectomies undertaken by surgeons during their career ( $\rho = 0.85$ ,  $P < 0.0001$ ). They declared mean work week duration of 66 hours. During the one year inclusion period, six (21%) surgeons did fewer than 30 thyroid procedures, seven (25%) did between 30 and 99 procedures, nine (32%) did between 100 and 199 procedures, and six (21%) did at least 200 procedures.

Univariate analysis showed that patients had permanent hypoparathyroidism more frequently after being operated on by surgeons who had spent the shortest or the longest time in practice since graduation (table 2). Postoperative hypoparathyroidism was associated with younger patients, female patients, and heavier thyroid specimens; Graves' disease or malignant neoplasm; and extended thyroidectomy or lymph node dissection. The frequency of complex surgical cases handled by surgeons was not associated with their experience. Surgeons in practice for at least 20 years did not have a higher proportion of high risk patients with Graves' disease, thyroid cancer, extended thyroidectomy, or lymph node dissection than their less experienced colleagues (21.1% (194/921) v 22.8% (315/1384) for nerve palsy analysis; 25.9% (286/1105) v 28.3% (476/1680) for hypoparathyroidism analysis).

Table 3 shows multivariable analyses of factors independently associated with complications after adjustment for both patient and surgeon variables. Experience of 20 years or more was the only factor associated with an increased probability of both recurrent laryngeal nerve palsy (odds ratio 3.06,  $P = 0.04$ ) and hypoparathyroidism (7.56,  $P = 0.01$ ). Occurrence of nerve palsy was linked with female patients (2.74,  $P = 0.03$ ), whereas hypoparathyroidism was associated with young patients (0.84 by 10 years increase,  $P = 0.04$ ) and inexperienced surgeons (5.93,  $P = 0.02$ ). Patients accounted for the largest proportion of total variability in surgical outcome. Variation attributable to surgeons was greater for hypoparathyroidism (32%) than for recurrent laryngeal nerve palsy (10%).

We found no association between recurrent laryngeal nerve palsy and a surgeon's length of experience and age (fig 3, web appendix). However, we saw a concave association between hypoparathyroidism and the length of experience or age of the surgeon ( $P = 0.036$  and  $P = 0.035$ , respectively; fig 3). Surgeons between 35 and 50 years old (that is, with 5-20 years of practice since graduation) had better outcomes than their younger and older colleagues.



## Discussion

In this multicentre study, patients were at increased risk of permanent complications after a thyroidectomy when operated on by inexperienced surgeons or those who have spent the longest time in practice since graduation. We observed a concave association between surgeons' experience and their case mix adjusted performance, suggesting that surgeons aged 35 to 50 years provided the safest care. Thyroidectomies by surgeons in practice for 20 years or more increased the probability of permanent complications considerably. Large effect sizes might raise important issues about patient safety and surgeons' experience, but those findings should be interpreted cautiously in the light of wide confidence intervals, owing to the scarcity of monitored complications.

## Comparison with other studies

Few studies have specifically assessed the influence of surgeons' experience on surgical outcomes. Older surgeons who have been in practice for a long time have been found to be associated with increased mortality rates after coronary artery bypass grafting<sup>12</sup> or carotid endarterectomy.<sup>13</sup> Increased experience, especially among doctors with more than 20 years' experience in practice,<sup>24</sup> might also be associated with a raised risk of inpatient death related to medical care.<sup>8</sup> However, compared with medical specialties, poor outcomes among very experienced surgeons could be explained by a weariness related to a high volume of procedures rather than improper technical skills or a lack of compliance with evidence based practice. Indeed, the organisation of healthcare provision in French academic hospitals is hierarchical, with most experienced surgeons at the head of surgical departments and responsible for the education and training of their young colleagues. Attending surgeons within the same team often reproduce similar techniques based on knowledge and skills that they have learned from older surgeons. Thyroid surgery is a highly reproducible and well defined procedure that has not changed substantially over the past decades.<sup>25</sup> Furthermore, most experienced surgeons generally spend more time on academic and administrative duties than their younger colleagues, who can focus more on surgical activities. The additional burden on workload related to non-clinical tasks could affect surgeons' attention in the operating room and jeopardise patient safety.<sup>26</sup>

According to the surgeons' performance curves, they are expected to reach peak performance in mid-career after their education ends and their techniques are well mastered. Following this period, and without appropriate training and ongoing challenges, a surgeon's performance could decline over time because of mental fatigue related to the repetition of a specific procedure over a long time. Physiological factors such as reduced stress with age<sup>27</sup> or habit might also lead to poor compliance with new techniques, and then to increased complication rates. Talent and experience are not enough to guarantee safe surgery if a surgeon does not possess the motivation and willingness to progress.<sup>28</sup> To maintain a high level of performance for the rest of their careers, surgeons should continuously strive to critically examine the care delivered.<sup>29</sup> Based on those assumptions, recertification of surgeons older than 50 years could focus on mental coaching and raising awareness to their own performance. Such an approach needs the implementation of outcome monitoring systems that have proved to be useful in reducing surgical complications in both local<sup>30</sup> and national<sup>31</sup> initiatives.

Previous analyses based on administrative databases have also established that surgeons with high volume workloads could

provide better care in thyroid surgery.<sup>32-34</sup> In this study, we did not identify any volume threshold or a particular time of day that was related to increased risks of complication. The daily operation schedule seemed to have marginal effects on the occurrence of complications, compared with surgeon experience and patient case mix variables. Similar to previous results, patient unit related factors accounted for the majority of variability in surgical outcome.<sup>35</sup> Irrespective of the lower proportion of variability attributable to surgeon and hospital centre levels, meaningful variation in performance could still be present even among a homogeneous population of highly specialised surgeons.

## Strengths and limitations of study

The strengths of our study included that it was designed a priori to model the nature of the association between surgeons' experience and performance in thyroid surgery, patient recruitment and data were recorded prospectively with great thoroughness, we based objective measurements of performance on rigorous follow-up of permanent complications that were systematically assessed and collected (and not self reported by surgeons), the patient case mix and other surgeon factors were partly controlled, and we considered the clustering of patients within surgeon and hospital centre levels.

However, this work had several limitations. The applicability of our results to other surgical fields is questionable, in view of the limited sample size of endocrine surgeons in academic referral centres. Our study had few middle aged surgeons with an intermediate length of experience, suggesting that the cohort of surgeons examined might not be representative of the average population of surgeons. The validity of our results also depends on how surgeons' experience was measured and whether the risk adjusted rates of complications showed surgical outcomes adequately. The length of experience was highly correlated with the surgeon's age or volume of thyroidectomies previously undertaken, but our analysis did not account for the possibility that a surgeon took a short career break, or that successive generation of surgeons might have had different experiences of training before entering independent practice.

Furthermore, we cannot exclude that other unknown or unmeasured factors might have explained part of the variation seen in complication rates. Despite adjusting for high risk patients and very complex surgical cases between the various surgeons' age categories, the singularity of thyroid diseases occasionally required procedures in which complexity might not have been sufficiently captured in our multivariate models. Surgical outcome depends on interactions between many factors that are still poorly understood and difficult to characterise.<sup>36</sup> These factors include the combination of manual and intellectual skills acquired during a surgeon's academic and professional training. The physical and mental condition of the surgeon on a given day is also essential, and a sudden or more insidious fatigue could reduce the surgeon's vigilance and increase the risk of complications.<sup>37-39</sup> The number of procedures undertaken,<sup>40-41</sup> excessive workloads,<sup>30-42</sup> or resident intraoperative participation<sup>43</sup> have also been suggested to influence surgical outcome. Although technical skills are a prerequisite for successful surgery, collective factors and a surgeon's leadership are essential for effective teamwork within the operating room.<sup>44-45</sup>

Future studies should be conducted with larger populations of surgeons in various settings and other surgical specialties to corroborate the potential link between experience and performance. Since a cross sectional study might be

inappropriate to resolve a dynamic phenomenon, a recommended design would be to follow a particular cohort of surgeons over time. Indeed, poor outcomes among older surgeons could be because their training is outdated rather than because of a declining performance with age. Despite practical constraints and possible confounding related to secular trends,<sup>8</sup> longitudinal outcome monitoring of surgeons can be implemented to explore changes in performance during their careers.

## Conclusions and policy implications

Working in a high volume academic hospital does not compensate for the probable lack of experience of newly appointed endocrine surgeons. Our findings also suggest that a surgeon cannot achieve or maintain top performance passively by accumulating experience, which raises concerns about ongoing training and motivation throughout a career that extends several decades. Solutions to help surgeons avoid poor outcomes could include simulation and proctoring in the early years of their careers, continuous monitoring of performance, and targeted retraining if appropriate. Individual feedback based on outcome indicators might increase awareness about performance and improve safety in surgical practice.

Members of the CATHY Study Group: Laurent Arnalsteen, Robert Caizzo, Bruno Carnaille, Guelareh Dezfoulian, Carole Eberle, Ziad El Khatib, Emmanuel Fernandez, Antoine Lamblin, François Pattou, and Marie-France Six (Lille); Stéphanie Bourdy, Laetitia Bouveret, Cyrille Colin, Antoine Duclos, Benoît Guibert, Marie-Annick Le Pogam, Jean-Christophe Lifante, Jean-Louis Peix, Gaétan Singier, Pietro Soardo, Sandrine Touzet, and Nicolas Voirin (Lyon); Pascal Auquier, Jean-François Henry, Claire Morando, Frédéric Sebag, and Sam Van Slycke (Marseille); Inès Akrouf, Fares Benmiloud, Jean-Paul Chigot, Isabelle Colombet, Gaëlle Godiris-Petit, Pierre Leyre, Fabrice Ménégau, Séverine Noullet, Benoît Royer, and Christophe Tresallet (Paris); Thibault Desurmont, Claudia Dominguez, Jean-Louis Kraimps, Chiara Odasso, and Laetitia Rouleau (Poitiers); Yves-Louis Chapuis, Pierre Durieux, Alain Lepape, and Frédéric Triponez (scientific committee). We thank Kathy Corso (Boston) for reviewing the English language in the paper.

Contributors: AD and JCL obtained funding and supervised the study. AD, CC, FM, FP, FS, JCL, JLK, JLP, and ST conceived and designed the study. CC, FM, FP, FS, JCL, SB, JLK, and JLP were responsible for data acquisition and provided administrative, technical, or material support. AD, JCL, NV, and SB analysed and interpreted the data. AD, JCL, and NV drafted the manuscript. AD, CC, FM, FP, FS, JCL, JLK, JLP, NV, SB, and ST revised the manuscript critically. All authors approved the final version of the manuscript to be published. AD is guarantor. All authors had full access to the data and take responsibility for its integrity and the accuracy of the analysis.

Funding: This study was supported by a grant from the Programme de Recherche en Qualité Hospitalière 2007 of the French Ministry of Health (Ministère chargé de la Santé, Direction de l'Hospitalisation et de l'Organisation des Soins), Hospices Civils de Lyon, Lyon. The funding source had no involvement in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. Researchers were independent from the funder.

Competing interests: All authors have completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author) and declare: that all authors received support from the French Ministry of Health for the submitted work; no relationships with any company that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: This study was approved by the Research Committee for the Protection of Persons and the National Advisory Committee on Information Processing in Material Research in the Field of Health in accordance with ethical directives, in France.

Patient consent: Informed consent was obtained from participating surgeons and all patients for sharing their data.

Data sharing: No additional data available.

- Weiser TG, Regenbogen SE, Thompson KD, Haynes AB, Lipsitz SR, Berry WR, et al. An estimation of the global volume of surgery: a modelling strategy based on available data. *Lancet* 2008;372:139-44.
- Ergina PL, Cook JA, Blazeby JM, Boutron I, Clavien PA, Reeves BC, et al. Challenges in evaluating surgical innovation. *Lancet* 2009;374:1097-104.
- De Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care* 2008;17:216-23.
- Ghaferi AA, Birkmeyer JD, Dimick JB. Variation in hospital mortality associated with inpatient surgery. *N Engl J Med* 2009;361:1368-75.
- Dionigi G, Bacuzzi A, Bertocchi V, Carratiello G, Boni L, Rovera F, et al. Safe incorporation of new technologies in thyroid surgery. *Expert Rev Med Devices* 2008;5:747-58.
- Barbash GI, Glied SA. New technology and health care costs—the case of robot-assisted surgery. *N Engl J Med* 2010;363:701-4.
- Ericsson KA. Deliberate practice and acquisition of expert performance: a general overview. *Acad Emerg Med* 2008;15:988-94.
- Choudhry NK, Fletcher RH, Soumerai SB. Systematic review: the relationship between clinical experience and quality of health care. *Ann Intern Med* 2005;142:260-73.
- Reznick RK, MacRae H. Teaching surgical skills—changes in the wind. *N Engl J Med* 2006;355:2664-9.
- Ramsay CR, Grant AM, Wallace SA, Garthwaite PH, Monk AF, Russell IT. Statistical assessment of the learning curves of health technologies. *Health Technol Assess* 2001;5:1-79.
- Carty MJ, Chan R, Huckman R, Snow D, Orgill DP. A detailed analysis of the reduction mammoplasty learning curve: a statistical process model for approaching surgical performance improvement. *Plast Reconstr Surg* 2009;124:706-14.
- Hartz AJ, Kuhn EM, Pulido J. Prestige of training programs and experience of bypass surgeons as factors in adjusted patient mortality rates. *Med Care* 1999;37:93-103.
- O'Neill L, Lanska DJ, Hartz A. Surgeon characteristics associated with mortality and morbidity following carotid endarterectomy. *Neurology* 2000;55:773-81.
- Munch S, de Kryger L. A piece of my mind. Moral wounds: complicated complications. *JAMA* 2001;285:1131-2.
- British Thyroid Association and Royal College of Physicians. Guidelines for the management of thyroid cancer. 2nd ed. British Thyroid Association and RCP, 2007.
- Agence Nationale d'Accréditation et d'Évaluation en Santé. Clinical practice guidelines. Recurrent laryngeal nerve paralysis in adults. *Ann Otolaryngol Chir Cervicofac* 2003;120:182-99.
- Jeannon JP, Orabi AA, Bruch GA, Abdalsalam HA, Simo R. Diagnosis of recurrent laryngeal nerve palsy after thyroidectomy: a systematic review. *Int J Clin Pract* 2009;63:624-9.
- Pattou F, Combemale F, Fabre S, Carnaille B, Decouls M, Wemeau JL, et al. Hypocalcemia following thyroid surgery: incidence and prediction of outcome. *World J Surg* 1998;22:718-24.
- Direction de la Recherche, des Etudes, de l'Évaluation et des Statistiques. Statistique annuelle des établissements de santé. 2008. [www.sae-diffusion.sante.gouv.fr/Collecte\\_2008/dwd\\_dwsen3.aspx?Acp=20101127](http://www.sae-diffusion.sante.gouv.fr/Collecte_2008/dwd_dwsen3.aspx?Acp=20101127).
- Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. *J Am Stat Assoc* 1993;88:9-25.
- R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, 2009.
- Brookhart MA, Solomon DH, Wang P, Glynn RJ, Avorn J, Schneeweiss S. Explained variation in a model of therapeutic decision making is partitioned across patient, physician, and clinic factors. *J Clin Epidemiol* 2006;59:18-25.
- Schemper M. Predictive accuracy and explained variation. *Stat Med* 2003;22:2299-308.
- Southern WN, Bellin EY, Arnsten JH. Longer lengths of stay and higher risk of mortality among inpatients of physicians with more years in practice. *Am J Med* 2011;124:868-74.
- Hegner CF. A history of thyroid surgery. *Ann Surg* 1932;95:481-92.
- Terry R, Hussain A, Nelson MF. Changing the "Peter Principled" culture of health care systems. *Physician Exec* 2011;37:26-8, 30.
- Stone AA, Schwartz JE, Broderick JE, Deaton A. A snapshot of the age distribution of psychological well-being in the United States. *Proc Natl Acad Sci USA* 2010;107:9985-90.
- Wallace JE, Lemaire JB, Ghali WA. Physician wellness: a missing quality indicator. *Lancet* 2009;374:1714-21.
- Duclos A, Lifante JC, Ducarroz S, Soardo P, Colin C, Peix JL. Influence of intraoperative neuromonitoring on surgeons' technique during thyroidectomy. *World J Surg* 2011;35:773-8.
- Duclos A, Touzet S, Soardo P, Colin C, Peix JL, Lifante JC. Quality monitoring in thyroid surgery using the Shewhart control chart. *Brit J Surg* 2009;96:171-4.
- Shahian DM, Edwards FH, Jacobs JP, Prager RL, Normand SL, Shewan CM, et al. Public reporting of cardiac surgery performance: part 1—history, rationale, consequences. *Ann Thorac Surg* 2011;92(suppl 3):S2-11.
- Sosa JA, Bowman HM, Tielsch JM, Powe NR, Gordon TA, Udelsman R. The importance of surgeon experience for clinical and economic outcomes from thyroidectomy. *Ann Surg* 1998;228:320-30.
- Stavrakis AI, Iuarte PH, Ko CY, Yeh MW. Surgeon volume as a predictor of outcomes in inpatient and outpatient endocrine surgery. *Surgery* 2007;142:887-99.
- Lifante JC, Duclos A, Couray-Targe S, Colin C, Peix JL, Schott AM. Hospital volume influences the choice of operation for thyroid cancer. *Br J Surg* 2009;96:1284-8.
- Fung V, Schmittluel JA, Fireman B, Meer A, Thomas S, Smider N. Meaningful variation in performance: a systematic literature review. *Med Care* 2010;48:140-8.
- Vincent C, Moorthy K, Sarker SK, Chang A, Darzi AW. Systems approaches to surgical quality and safety: from concept to measurement. *Ann Surg* 2004;239:475-82.

**What is already known on this topic**

Experts reach their peak performance in various specialties between the ages of 30 and 50 years

Doctors who have been in practice for a long time since graduation possess less factual knowledge and are less likely to adhere to evidence based care

Rates of postoperative complications can vary greatly during a surgeon's career

**What this study adds**

Patients were at an increased risk of permanent complications after a thyroidectomy by inexperienced surgeons and by those in practice for 20 years or more

Surgeons aged 35 to 50 years could have better postoperative outcomes than their younger and older colleagues, which raises concerns about ongoing training and motivation of surgeons during their careers

Variability in performance over the course of a surgeon's career are caused by unknown factors that should be further explored

- 37 Shanafelt TD, Balch CM, Bechamps G, Russell T, Dyrbye L, Satele D. Burnout and medical errors among American surgeons. *Ann Surg* 2010;251:995-1000.
- 38 Rothschild JM, Keohane CA, Rogers S, Gardner R, Lipsitz SR, Salzberg CA. Risks of complications by attending physicians after performing nighttime procedures. *JAMA* 2009;302:1565-72.
- 39 Pettiker CM, Funai EF. Getting it right when things go wrong. *JAMA* 2010;303:977-8.
- 40 Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. *N Engl J Med* 1979;301:1364-9.
- 41 Eggli Y, Halfon P, Meylan D, Taffé P. Surgical safety and hospital volume across a wide range of interventions. *Med Care* 2010;48:962-71.
- 42 Volpp KG, Landrigan CP. Building physician work hour regulations from first principles and best evidence. *JAMA* 2008;300:1197-9.
- 43 Raval MV, Wang X, Cohen ME, Ingraham AM, Bentrem DJ, Dimick JB, et al. The influence of resident involvement on surgical outcomes. *J Am Coll Surg* 2011;212:889-98.
- 44 Neily J, Mills PD, Young-Xu Y, Carney BT, West P, Berger DH, et al. Association between implementation of a medical team training program and surgical mortality. *JAMA* 2010;304:1693-700.
- 45 Catchpole K, Mishra A, Handa A, McCulloch P. Teamwork and error in the operating room: analysis of skills and roles. *Ann Surg* 2008;247:699-706.

**Accepted:** 20 October 2011

Cite this as: [BMJ 2012;344:d8041](https://doi.org/10.1136/bmj.d8041)

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license. See: <http://creativecommons.org/licenses/by-nc/2.0/> and <http://creativecommons.org/licenses/by-nc/2.0/legalcode>.

## Tables

**Table 1 | Characteristics of participating surgeons and hospital centres**

	Hospital centre					
	A	B	C	D	E	All
<b>Surgeon level</b>						
Participating surgeons	5	3	8	4	8	28
Female/male surgeon	0/5	0/3	1/7	2/2	2/6	5/23
Age (years)	38 (33-63)	38 (33-63)	35 (30-54)	32.5 (31-52)	34.5 (32-69)	35.5 (30-69)
Length of experience (years)*	7 (3-36)	7 (3-35)	3 (0-25)	1 (0-25)	3 (0-38)	3 (0-38)
Work week duration (hours)†	60 (50-80)	65 (60-70)	60 (50-92)	62.5 (40-66)	80 (50-90)	65 (40-92)
No of operated patients	101 (4-406)	239 (191-331)	62 (10-172)	90 (51-272)	133 (19-387)	109 (4-406)
<b>Hospital centre level</b>						
Acute care beds (surgical beds) in hospital	751 (290)	719 (327)	446 (266)	984 (320)	1459 (563)	4359 (1766)
Healthcare workers in hospital	3504	3571	1534	4792	8509	21910
Operating rooms in hospital	30	26	26	27	45	154
Patients included in study	744	747	560	501	1022	3574

Data are number or median values (minimum-maximum) unless stated otherwise.

\*Calculated as no of years spent in practice since graduation (that is, surgeons at the beginning of their career during the study period had a length of experience equal to zero).

†Declared by each surgeon based on the average number of hours spent weekly for surgical, administrative, teaching, and research activities.

Table 2| Distribution of surgeon and patient characteristics regarding permanent complications after thyroid surgery (univariate analysis)

Variable	No (%) of thyroid procedures					
	Recurrent laryngeal nerve palsy			Hypoparathyroidism		
	Yes (n=49)	No (n=2308)	P value	Yes (n=78)	No (n=2826)	P value
<b>Length of surgeon experience (years)*</b>						
<2 (n=8 surgeons)	8 (2.4)	323 (97.6)	0.06	13 (2.7)	477 (97.3)	0.02
2-4 (n=7 surgeons)	10 (1.7)	578 (98.3)		23 (3.0)	749 (97.0)	
5-19 (n=6 surgeons)	4 (0.8)	489 (99.2)		3 (0.6)	484 (99.4)	
≥20 (n=7 surgeons)	27 (2.9)	918 (97.1)		39 (3.4)	1116 (96.6)	
<b>Surgical procedures performed by surgeon on same day (n)</b>						
<3	16 (2.6)	588 (97.4)	0.53	24 (3.2)	726 (96.8)	0.67
3-4	26 (1.9)	1372 (98.1)		47 (2.6)	1744 (97.4)	
≥5	6 (2.2)	273 (97.8)		6 (2.4)	248 (97.6)	
<b>Thyroid disease</b>						
Non-toxic solitary nodule	3 (0.9)	340 (99.1)	0.24	0	121 (100)	0.04
Non-toxic multinodular goitre	27 (2.1)	1275 (97.9)		44 (2.4)	1755 (97.6)	
Hyperthyroidism	7 (2.8)	246 (97.2)		6 (2.1)	286 (97.9)	
Graves' disease	8 (3.5)	220 (96.5)		11 (3.4)	313 (96.6)	
Malignant neoplasm	4 (1.7)	227 (98.3)		17 (4.6)	351 (95.4)	
<b>Thyroid procedure</b>						
Unilateral lobectomy	4 (1.2)	336 (98.8)	0.51	Not included	Not included	0.02
Subtotal thyroidectomy	1 (5.6)	17 (94.4)		0	24 (100)	
Total thyroidectomy	41 (2.2)	1847 (97.8)		69 (2.6)	2600 (97.4)	
Extended thyroidectomy	1 (4.8)	20 (95.2)		4 (10.3)	35 (89.7)	
Completion thyroidectomy	2 (2.2)	88 (97.8)		5 (2.9)	167 (97.1)	
<b>Lymph node dissection</b>						
Yes	2 (1.0)	189 (99.0)	0.29	15 (4.5)	318 (95.5)	0.03
No	47 (2.2)	2104 (97.8)		60 (2.4)	2414 (97.6)	
<b>Patient sex</b>						
Male	6 (1.1)	524 (98.9)	0.08	9 (1.4)	617 (98.6)	0.03
Female	43 (2.4)	1784 (97.6)		69 (3.0)	2208 (97.0)	
<b>Patient age (years)</b>						
Mean (SD)	51.5 (15.0)	51.3 (14.9)	0.95	48.1 (16.4)	51.8 (14.6)	0.02
<b>Patient body mass index†</b>						
Mean (SD)	26.5 (6.9)	25.8 (5.2)	0.88	25.7 (5.8)	25.9 (5.2)	0.46
<b>Weight of thyroid specimen (g)</b>						
Mean (SD)	49.3 (59.5)	45.8 (48.2)	0.56	51.1 (40.7)	48.2 (48.9)	0.05

\*Calculated as the number of years spent in practice since graduation.

†Calculated as weight (kg) divided by height (m<sup>2</sup>).



Table 3| Factors independently associated with occurrence of permanent complications after thyroid surgery (multivariate analysis)

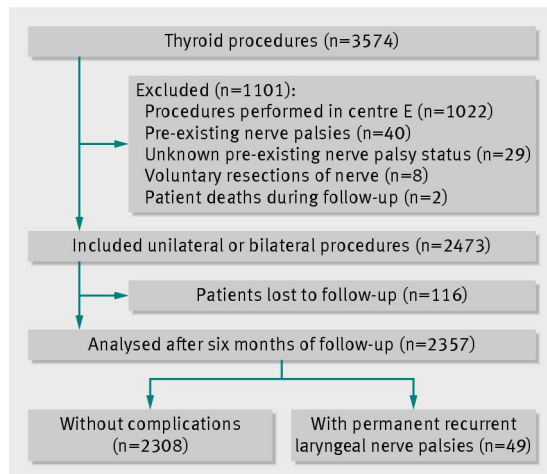
Variable	Permanent complications (odds ratio (95% CI))			
	Recurrent laryngeal nerve palsy		Hypoparathyroidism	
	Unadjusted	Adjusted†	Unadjusted	Adjusted‡
<b>Length of surgeon experience</b>				
<2 years	3.59 (0.98 to 13.10)	2.41 (0.69 to 8.36)	4.44 (1.15 to 17.10)	5.93 (1.28 to 27.39)
2-4 years	2.34 (0.64 to 8.53)	1.77 (0.55 to 5.69)	4.77 (1.33 to 17.16)	5.80 (1.34 to 25.10)
5-19 years	1.00	1.00	1.00	1.00
≥20 years	2.78 (0.83 to 9.32)	3.06 (1.07 to 8.80)	5.24 (1.50 to 18.34)	7.56 (1.79 to 31.99)
<b>Surgical procedures performed by surgeon on same day (n)</b>				
<3 procedures	1.00	1.00	1.00	1.00
3-4 procedures	0.76 (0.42 to 1.37)	0.90 (0.46 to 1.79)	0.73 (0.45 to 1.20)	0.84 (0.49 to 1.46)
≥5 procedures	0.84 (0.33 to 2.12)	1.15 (0.42 to 3.20)	0.65 (0.26 to 1.58)	0.73 (0.26 to 2.05)
<b>Type of surgical case</b>				
Routine	1.00	1.00	1.00	1.00
Complex*	1.24 (0.69 to 2.24)	1.56 (0.80 to 3.07)	1.58 (1.00 to 2.51)	1.51 (0.89 to 2.55)
<b>Patient sex</b>				
Male	1.00	1.00	1.00	1.00
Female	2.25 (1.02 to 4.98)	2.74 (1.08 to 6.97)	2.15 (1.10 to 4.22)	1.98 (0.96 to 4.06)
<b>Patient age</b>				
By 10 year increase	1.00 (0.98 to 1.02)	1.02 (0.83 to 1.25)	0.98 (0.97 to 0.99)	0.84 (0.71 to 0.99)
<b>Patient body mass index</b>				
By 5 point increase	1.04 (0.99 to 1.09)	1.21 (0.93 to 1.57)	0.99 (0.95 to 1.04)	1.06 (0.84 to 1.33)
<b>Weight of thyroid specimen</b>				
By 100 g increase	1.00 (0.99 to 1.01)	1.23 (0.71 to 2.15)	1.00 (0.99 to 1.01)	1.34 (0.83 to 2.14)

\*Included a surgical indication of Graves' disease or malignant neoplasm, an extended thyroidectomy, or a thyroidectomy with lymph node dissection.

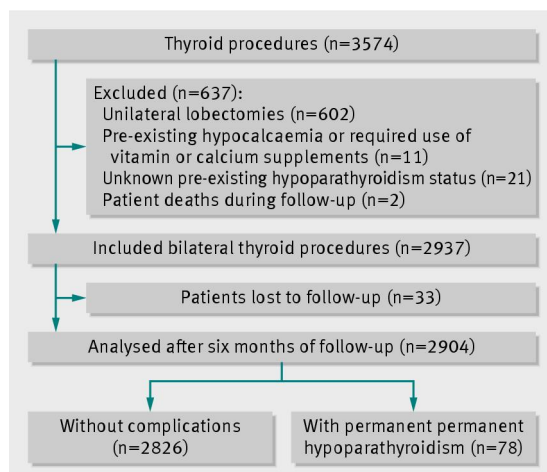
†Variation attributable to the patient, surgeon, and hospital centre levels was 58.6%, 10.3%, and 0.0%, respectively.

‡Variation attributable to the patient, surgeon, and hospital centre levels was 55.6%, 31.5%, and 0.0%, respectively.

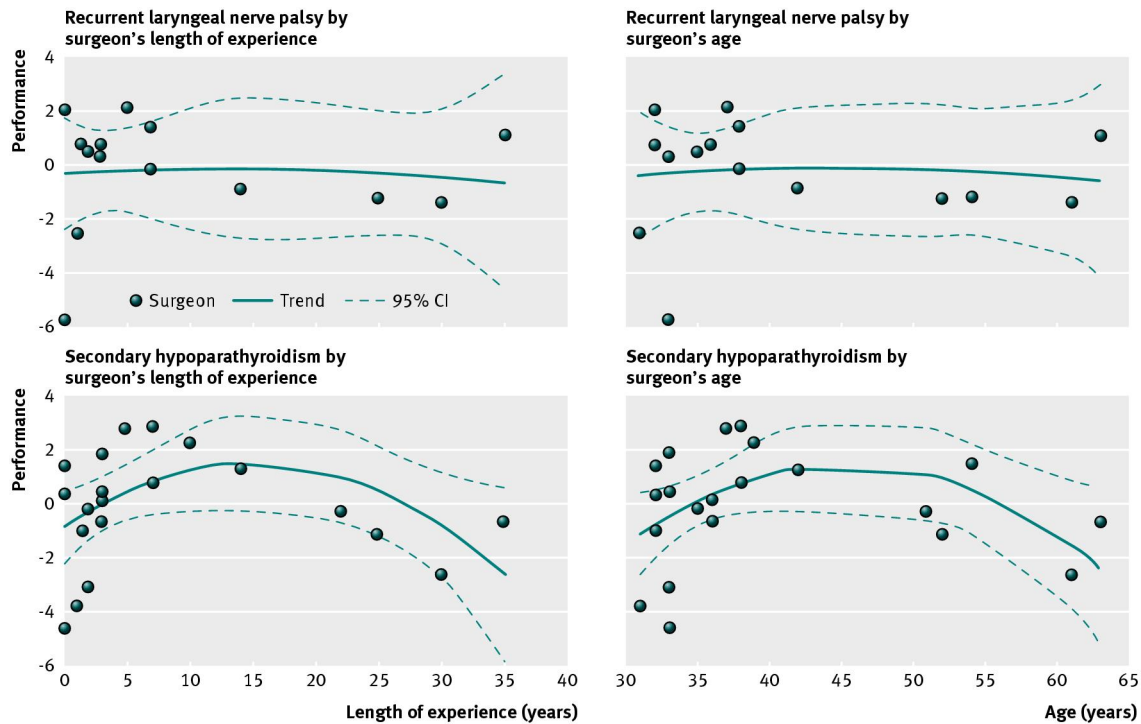
## Figures



**Fig 1** Flow diagram of procedures in analysis of recurrent laryngeal nerve palsy



**Fig 2** Flow diagram of procedures in analysis of hypoparathyroidism



**Fig 3** Surgeon performance curve by permanent complication. Points were above zero if the surgeon's performance was higher than predicted by the risk model, and below zero if the performance was lower than predicted.