

# Spreadsheet Error Types and Their Prevalence in a Healthcare Context

Elaine Dobell, Saolta University Health Care Group, Galway, Ireland & University of Limerick, Limerick, Ireland

Sebastian Herold, Department of Mathematics and Computer Science, Karlstad University, Karlstad, Sweden

Jim Buckley, CSIS / Lero – The Irish Software Research Centre, University of Limerick, Limerick, Ireland

## ABSTRACT

Spreadsheets are commonly used to inform decision making across many business sectors, despite the fact that research performed in the financial sector has shown that they are quite error-prone. However, few studies have investigated spreadsheet errors and their impact in other domains, like the healthcare sector. This article derives a lifecycle-stage classification scheme of spreadsheet error types based on an aggregation of, and extension of, existing classifications. Based on these classifications, a case study is then presented, performed to investigate the prevalence of these spreadsheet error types in an Irish healthcare setting. Results reveal that more than 90% of the spreadsheets studied contained ‘bottom-line’ errors and the average cell-error rate was 13%. There was also a correlation between increased perceived impact of the spreadsheets and the number of errors identified. Recommendations from this research include providing spreadsheet training and guidelines for developers and users, and systematically managing and auditing spreadsheet development and use.

## KEYWORDS

Auditing, Case Study, Classification, Error Types, Errors, Healthcare, Spreadsheet, Taxonomy

## INTRODUCTION

Spreadsheets usage is widespread across all sectors of commercial organizations (Pemberton and Robson 2000, Kruck 2006), most notably in the financial sector (Madininos et al. 2012, Burdick 2008, Ayalew et al. 2000, Read & Batson 1999, Baxter 2012). Not only are spreadsheets prevalent but they are also commonly of high impact, being used as a key tool in organizations’ operational activities and financial reporting (PWC 2004). Kruck (2006), for example, states that “business professionals use spreadsheets extensively” and that “decisions made using spreadsheets involve billions of dollars”. Their usage in the healthcare sector is less reported, but early work by Croll and Butler (2006) suggests that spreadsheets are prevalent in that domain also.

However, spreadsheet development and usage is error-prone with spreadsheet errors considered “common and non-trivial” (Panko 2000). Reviews of the literature in Kruck (2006) and Panko (1998), show that the numbers of operational and laboratory spreadsheets with errors range between 7% and 82%, with an average of 40% of professional spreadsheets containing errors. KPMG, an international auditing company, (cited in Rajalingham et al. 2000), found over five errors in 95% of the financial models they reviewed. Likewise, Coopers and Lybrand found that 90% of the spreadsheets they audited had errors (Rajalingham et al. 2000) and the “Tuck School of Business at Dartmouth found that 94% of spreadsheets and 5% of all formulae within spreadsheets contain errors” (Arnott 2011).

DOI: 10.4018/JOEUC.2018040102

Copyright © 2018, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

One of the most frequently mentioned reasons for the high error rate in spreadsheets, is the fact that spreadsheets are developed by end-users, having little formal education in computing and lacking knowledge of structured design methods (Ahmad et al. 2003, Powell et al. 2009a).

Information about the actual impact of spreadsheet error is scarce, but severe when reported: Powell et al. (2009a) reported on several instances of erroneous spreadsheets in the financial sector having stark financial impact, with one organisation suffering a \$100 million impact. Additionally, the EuSpRIG website reports many anecdotal examples of very poor outcomes due to spreadsheet error (O’Beirne et al. 2013). But overall, there is little academic literature reporting of the risk or impact of these errors, probably due to the sensitivity of the information. Regardless, it can be assumed, that spreadsheets are not used as efficiently and reliably as they could be.

Given the reported prevalence and probable low spreadsheet quality across sectors, the focus of this research has been to investigate error in spreadsheets. The decision to focus on the Healthcare sector arises from the reported prevalence of spreadsheet use in this domain Croll & Butler (2006), and the potentially serious consequences of errors potentially introduced by spreadsheet development and use in this domain: incorrect or poor clinical decisions (resulting in patient harm), loss of finances or resources for appropriate healthcare agendas, or reputational harm. It is therefore important to identify the prevalence of various error types in healthcare spreadsheets so that the problem can be quantified and remedial actions taken to reduce the error rate. Healthcare, in the context of this research, refers to acute hospitals delivering frontline care. Croll & Butler (2006) demonstrated significant errors in three open spreadsheets used for clinical decision making and service planning which could have resulted in significant concerns for patient safety. But apart from this one study, our literature review found little other research into spreadsheet error with respect to these two concerns.

This research moves towards addressing this lack of empirical evidence, focusing on the service planning aspect of healthcare and specifically addressing the following research questions:

- How prevalent is spreadsheet error in healthcare administrative spreadsheets?
- What types of errors occur in these healthcare spreadsheets?

In order to accurately answer these questions, a taxonomy of existing spreadsheet error-types is a pre-requisite (Purser & Chadwick 2008). Hence, a taxonomy was aggregated from the literature and coalesced into a single taxonomy. This taxonomy was then used to identify the errors in healthcare-administrative spreadsheets in a case-study performed with an Irish hospital. Twelve different spreadsheets were investigated and audited for errors.

The remaining article is structured as follows. First, it gives an overview of the state of the art of spreadsheet error categorization schemes. Then, the design of the study will be presented, followed by the results of the study, a discussion of the results and our conclusions.

## STATE OF THE ART

In this section, relevant research on taxonomies of spreadsheet errors is reviewed. The search strategy employed to identify the taxonomy papers in this area was to source them through academic (Google scholar) and general web search engines (Google), using the term ‘spreadsheet error’. Subsequently, forward and backward reference-chaining was used on the results. The papers were then filtered for relevance with respect to spreadsheet error-taxonomies and protocols for detecting spreadsheet errors. This filtering was based on their title, abstract and in several ambiguous cases, a more complete review. The resultant papers are noted in column one of Table 1.

An initial review of these papers suggested that the taxonomies also regard poor developer practices, which may cause subsequent faults in spreadsheets, as errors worthy of inclusion. Hence an additional search was performed on the same search engines for ‘spreadsheet best practices’, in

21 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/article/spreadsheet-error-types-and-their-prevalence-in-a-healthcare-context/197349](http://www.igi-global.com/article/spreadsheet-error-types-and-their-prevalence-in-a-healthcare-context/197349)

## Related Content

---

### Understanding the E-Banking Channel Selection Behavior of Elderly Customers: A Small-World Network Perspective

Xiaolei Cui and Jianzhong Xu (2022). *Journal of Organizational and End User Computing* (pp. 1-21).

[www.irma-international.org/article/understanding-the-e-banking-channel-selection-behavior-of-elderly-customers/300765](http://www.irma-international.org/article/understanding-the-e-banking-channel-selection-behavior-of-elderly-customers/300765)

### Overreliance on Mathematical Accuracy of Computer Output: An Issue for IT Educators

Ian Sims, Conor O'Leary and Pran Boolaky (2014). *Journal of Organizational and End User Computing* (pp. 47-64).

[www.irma-international.org/article/overreliance-on-mathematical-accuracy-of-computer-output-an-issue-for-it-educators/116695](http://www.irma-international.org/article/overreliance-on-mathematical-accuracy-of-computer-output-an-issue-for-it-educators/116695)

### Marketing Decision Model and Consumer Behavior Prediction With Deep Learning

Anfeng Xu, Yue Li and Praveen Kumar Donta (2024). *Journal of Organizational and End User Computing* (pp. 1-25).

[www.irma-international.org/article/marketing-decision-model-and-consumer-behavior-prediction-with-deep-learning/336547](http://www.irma-international.org/article/marketing-decision-model-and-consumer-behavior-prediction-with-deep-learning/336547)

### e-Search: A Conceptual Framework of Online Consumer Behavior

Kuan-Pin Chiang, Ruby Roy Dholakia and Stu Westin (2008). *End-User Computing: Concepts, Methodologies, Tools, and Applications* (pp. 509-524).

[www.irma-international.org/chapter/search-conceptual-framework-online-consumer/18205](http://www.irma-international.org/chapter/search-conceptual-framework-online-consumer/18205)

### Application of Artificial Intelligence in Precision Marketing

Xue Yang, Haowen Li, Likun Ni and Teng Li (2021). *Journal of Organizational and End User Computing* (pp. 209-219).

[www.irma-international.org/article/application-of-artificial-intelligence-in-precision-marketing/280496](http://www.irma-international.org/article/application-of-artificial-intelligence-in-precision-marketing/280496)